Autonomous science for an ExoMars Rover-like mission

Woods, Mark; Shaw, Andy; Barnes, Dave; Price, David Ernest; Long, Derek; Pullan, Derek

Published in:
Journal of Field Robotics

DOI:
10.1002/rob.20289

Publication date:
2009

Citation for published version (APA):
Woods, M., Shaw, A., Barnes, D., Price, D. E., Long, D., & Pullan, D. (2009). Autonomous science for an ExoMars Rover-like mission. *Journal of Field Robotics*, *26*(4), 358-390. https://doi.org/10.1002/rob.20289

Autonomous Science for an ExoMars Rover–Like Mission

Mark Woods and Andy Shaw
SciSys UK Ltd.
Bristol BS4 5SS
United Kingdom
e-mail: mark.woods@scisys.co.uk, andy.shaw@scisys.co.uk

Dave Barnes and Dave Price
Department of Computer Science
Aberystwyth University
Aberystwyth SY23 3DB
United Kingdom
e-mail: dpb@aber.ac.uk, dap@aber.ac.uk

Derek Long
Department of Computer and Information Sciences
University of Strathclyde
Glasgow G1 1XH
United Kingdom
e-mail: derek.long@cis.strath.ac.uk

Derek Pullan
Department of Physics and Astronomy
University of Leicester
Leicester LE1 7RH
United Kingdom
e-mail: dpu@star.le.ac.uk

Received 2 July 2008; accepted 15 January 2009

In common with other Mars exploration missions, human supervision of Europe's ExoMars Rover will be mostly indirect via orbital relay spacecraft and thus far from immediate. The gap between issuing commands and witnessing the results of the consequent rover actions will typically be on the order of several hours or even sols. In addition,

Journal of Field Robotics 26(4), 358–390 (2009)
Published online in Wiley InterScience (www.interscience.wiley.com). • DOI: 10.1002/rob.20289

it will not be possible to observe the external environment at the time of action execution. This lengthens the time required to carry out scientific exploration and limits the mission's ability to respond quickly to favorable science events. To increase potential science return for such missions, it will be necessary to deploy autonomous systems that include science target selection and active data acquisition. In this work, we have developed and integrated technologies that we explored in previous studies and used the resulting test bed to demonstrate an autonomous, opportunistic science concept on a representative robotic platform. In addition to progressing the system design approach and individual autonomy components, we have introduced a methodology for autonomous science assessment based on terrestrial field science practice.

1. INTRODUCTION

The forthcoming ExoMars Rover mission scheduled for launch in 2016 is the European Space Agency's (ESA) first attempt to deploy a mobile robotic platform on the surface of Mars. The mission has an exobiology focus and will be used by scientists to search for signs of extinct or extant life on Mars (Van Winnendael, Baglioni, & Vago, 2005). The six-wheeled mobile platform will weigh on the order of 200 kg (Figure 1) and will have a dedicated science payload known as Pasteur, which will incorporate a range of panoramic, contact, and analytical laboratory instruments. This will include a drill to gather subsurface samples, ground-penetrating radar, wide-angle and close-up imaging capabilities, and a robotic arm to deploy contact instruments. The nominal 180-sol (a Martian solar day equal to 24 h, 39 min) science plan for ExoMars is to visit and investigate a number of primary sites that are on the order of 500 m apart and gather samples for analysis through successive and iterative sample assessment. The same basic process will apply at each site, i.e., broad site assessment, multiple target selection for deeper analysis, target assessment, continued refinement of the target selection, and analysis through the range of position and imaging scales leading to eventual sample selection and characterization.

Figure 1. ESA's ExoMars rover deployed on the surface: artist's impression.

At the time of writing it is expected that communication to the rover will be established mainly via a NASA Orbiter, and possibly a low-bandwidth direct-to-earth (DTE)/direct-from-earth (DFE) link will also be available. Consequently, there will be a lag between data acquisition and delivery to the science and engineering teams for assessment, which will inevitably introduce delays in the operations process. The volumes of data required for science assessment and planning are significant and will strain the available bandwidth. It may therefore take a number of link sessions to deliver all the data required for assessment and planning, which will lengthen the duration of the assessment and response cycle. The nominal science operations plan described above has a number of discrete ground decision points (GDPs) at which the science team must be involved, e.g., in selecting a small number of candidate targets in a wide-angle camera (WAC) image for further analysis. This entire planning cycle of acquire data, delayed delivery, assessment, planning response with delayed delivery of commands, execution, and assessment of command execution back on the ground is inherently inefficient. Maintaining this science plan will be a challenge given the disjoint nature of the communications path and the natural uncertainties associated with robot–Mars terrain interaction. As

the open-loop planning and response process is slow, any approach that can improve response times or reduce the need for Earth-based assessment is clearly attractive.

One way of trying to maintain the proposed science schedule is to consider migrating some of the assessment and planning tasks carried out by operations teams to the space segment in order to decrease decision-making response times. This of course implies greater onboard autonomy. By carrying out autonomous science assessment and response planning on the robotic platform itself, the bottleneck associated with ground-based decision making would be greatly reduced. However, it was clear prior to the commencement of this study that the technology required to support this concept is not sufficiently mature. It is too early therefore to consider replacing GDPs with an autonomous equivalent. There was, however, sufficient justification to evaluate the use of this approach for opportunistic field science, i.e., respond proactively to the immediate science environment without requiring the slower Earth-based assessment and response. Such an approach could also offer support in the form of lower order and automated data-gathering functions, particularly when anomalies occur. For example, an onboard assessment of acquired data could be used to quickly schedule a retry if the initial acquisition was corrupted or unsatisfactory for some reason. This could rectify a tactical planning problem in minutes rather than sols. The overall goal therefore is to develop an approach that would increase the science return per time spent on the surface for such a mission.

2. RATIONALE AND OBJECTIVES

The basic ExoMars science exploration model of repeated analysis at reasonably well-separated sites is well disposed toward opportunistic science, as Figure 2 shows. The terrain intervals between sites will be traversed "blind" from a (Earth-based) science perspective in order to maintain the primary science schedule. Of course images will be gathered (e.g., for navigation or in a preprogrammed way) en route, and these could be used to alter (by Earth-based ops teams) primary plans if a target of sufficient interest presents itself.

A significant enhancement would be to equip the platform with the ability to autonomously assess the terrain with respect to science goals and if necessary plan a response that provides deeper analysis of some candidate target. This would allow science to be obtained more selectively without having a great impact on the nominal site visit schedule.

Focusing on an opportunistic approach serves three purposes. First, the technology requirements for the science assessment that apply in this mode are less stringent than those associated with replacing the current GDPs as the intent is to improve on "unintelligent" data capture. The GDPs are on the critical path so the required classification performance must be excellent in order to maintain the mission schedule. Opportunistic acts are by contrast to be used in periods mainly when spare resources such as time and power are available. As long as tactical resource allocation and replanning functions are available onboard, the risk to the nominal operations is negligible as resource prioritization will be considered at the appropriate level. It is also configurable in the sense that autonomous responses to assessment decisions can be constrained by mission operators in advance.

Second, the solutions we develop for these problems will build a capability that could be used to augment or possibly replace the current set of GDPs over the longer term for future missions. Third, it is envisaged that low-level autonomous science selection and response replanning could play an important part in recovering from certain anomalies during the nominal science schedule and prioritizing data for download. It is envisaged that this work will help develop capability that supports this concept.

2.1. Related Work

Autonomous systems for planetary exploration are currently very topical and include studies in rock detection and target prioritization (Castaño et al., 2006, 2007; Pedersen, 2000), feature detection (Gulick, Morris, Ruzon, & Roush, 2001; Thompson, Smith, & Wettergreen, 2005), novelty detection (Thompson, Smith, & Wettergreen, 2006), compositional evaluation (Gazis & Roush, 2001), and life detection (Wettergreen et al., 2005). These studies often focus on isolated or float rocks as potential science targets with the aim of assigning basic scientific parameters to them such as albedo and basic texture and color and in some cases consider analytical input such as spectral signature and fluorescence indicators. Results can be used to detect targets of interest and unexpected objects and characterize the exploration site. Some employ techniques common to navigation and locomotion work in which landscape features such as

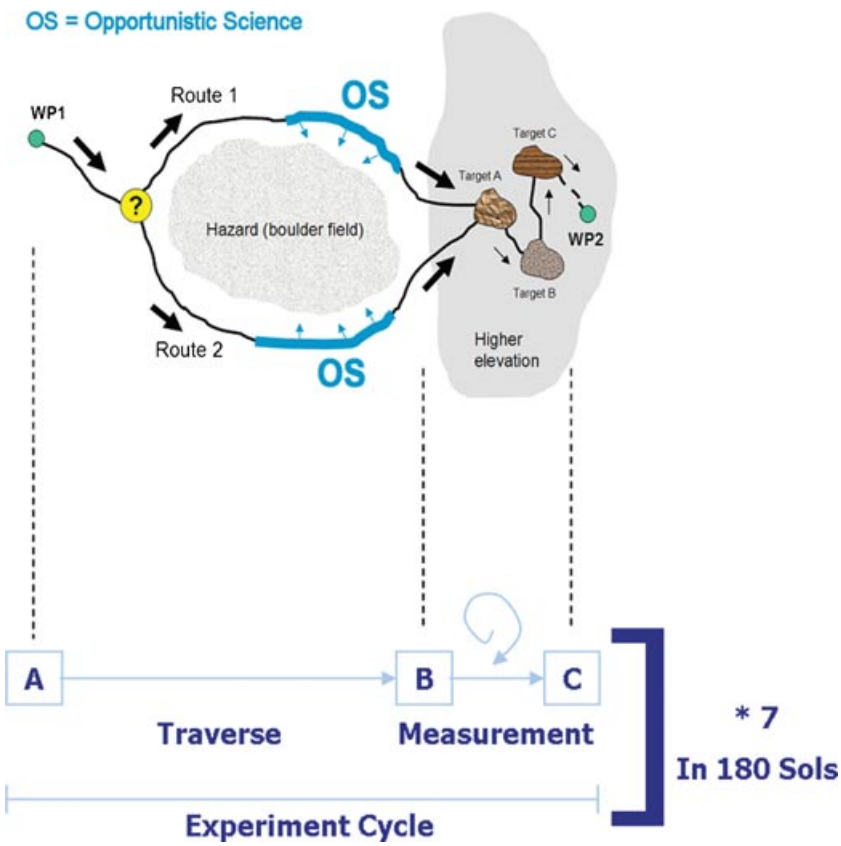

Figure 2. Outline of basic ExoMars exploration cycle showing the potential for opportunistic science activities. The main objective is to traverse from a previously explored site (A) and progress toward the next site at B, where detailed sample assessment will be carried out. The intention is to visit seven sites over the nominal 180-sol period. Opportunistic science is clearly possible during the traverse phase and could be used to improve the robustness of data acquisition and prioritization during the measurement cycle.

rocks and slopes are considered "obstacles," thereby requiring some detailed autonomous assessment of physical parameters such as shape and size.

Like those of many colleagues working in the discipline, our approach to scientific autonomy is based on fundamental human geological field practice. The evaluation of complex geological features is achieved by recognizing, classifying, and associating individual elements that make up the feature and reevaluating the assessment with scale.

A unique aspect of this work is that the geological remit is broader as we are interested in understanding how geologists actively bring wider knowledge to bear when identifying regions of interest in a complex scene. Autonomous target identification is pursued with the active (and autonomous) acquisition and analysis of additional data describing the target and host scene at different scales, reflecting techniques employed in human field practice. In common with other approaches, we also consider the need to incorporate mission objectives into our scientific assessments in order to influence what is more (or less) relevant at any particular mission phase or locality. Research by others in this area has sought to isolate or detect discrete targets of interest in unstructured and complex terrain. For example, the work of Pedersen (2000) focuses on discriminating meteorites from terrestrial rocks, the primary objective being to find meteorites. Our approach is intended to extend the geological scope and flexibility of the final system by considering what is important at the time (i.e., evidence of layering, hydrated minerals, fossilized

biosignatures, etc.) and incorporating this deeper knowledge into our assessment of the target area. This requires a broader geological "knowledge" and a broader sense of prioritization across a number of subdisciplines within geology. Detecting individual rocks remains important in target selection, but our focus extends to geological features of potentially higher value across the scale range from large stratigraphic sequences to microscopic signatures of life. This is supported by another unique aspect of our work, which is to incorporate instrument (and tool) placement and positioning into the autonomous science activity. This emulates the essential human traits of observation, interpretation, and physical interaction. In the long term, we are ultimately aiming to emulate a field exploration geologist who would perform at a higher level than a meteorite hunter or someone performing similarly focused tasks. An introduction to the methodology that will form the basis for future development and represents our first steps in this regard is presented in Section 4.1.

In addition to autonomous science target identification, our research has focused upon the problems associated with autonomous arm placement, i.e., given a target rock, the operation of moving and contacting an instrument against the science target using autonomous arm control methods. A large body of work has been undertaken in this area (Backes, Diaz-Calderon, Robinson, Bajracharya, & Helmick, 2005; Huntsberger, Cheng, Stroupe, & Aghazarian, 2005; Pedersen et al., 2006) and comes under the acronyms of SCIP (single-cycle instrument placement) or SCAIP (single command approach and instrument placement). The main drive for this work has been the desire to maximize the science data return rate by limiting the number of required command cycles for each individual instrument placement operation. An added benefit is to minimize the ground-based operator workload. Amortized primary mission NASA Mars Exploration Rover (MER) operations have been reported to cost approximately $4 million to $4.5 million per day and require 240 operators working 24/7 (Pedersen, Deans, Lees, Rajagoplan, & Smith, 2005). The SCIP goal is to autonomously approach and place an instrument on multiple features of scientific interest in a single command sequence uplink (Pedersen et al., 2006). Vision-based target tracking techniques are key to the general SCIP approach, and two-dimensional (2D) feature–based visual servoing has been used to keep a rover's navigation cameras foveated onto a science target and command the rover directly toward the given target. However, results have shown that whereas 2D appearance–based techniques can be computationally inexpensive, for visual servoing, they are insufficient as they can drift and are not robust to changes in a target's appearance as the rover's traverse is executed. Three-dimensional (3D) shape information is required, and stereo-based shape tracking techniques have been employed with good results (e.g., 10-m-distant target and up to 10-mm tracking and handoff accuracy). Such techniques are robust to image noise and changing lighting conditions, but they are sensitive to calibration parameters and can be computationally expensive. It should be noted that with the current SCIP work, it is a ground-based scientist(s) who identifies and selects the desired science target from a previously captured panoramic camera (pancam) image.

Our approach has been to integrate the desire for autonomous opportunistic science target identification with that of autonomous arm placement and hence lay the groundwork for autonomous science data acquisition. Our work therefore draws upon the results obtained from previous autonomous science and SCIP-based research. In particular we have based our autonomous opportunistic science target identification work on 2D appearance–based techniques because an opportunistic approach must be real time and cannot be computationally expensive. However, our autonomous arm placement work has been based on stereo-based techniques and 3D shape information. Although such techniques can be computationally expensive, we argue that this would be acceptable (within reason) given the clear benefits when actual science target contact is required and given the fact that the rover would not be executing a traverse with all the inherent additional demands of real-time collision detection and avoidance.

The main aim of our work was to prototype methods that are applicable to an ExoMars-like operations scenario, which has its own unique attributes and constraints, and to contribute to the long-term development of science autonomy techniques. The objective was not to implement flight-ready autonomous science for the actual ExoMars mission. As such it was part of the U.K. STFC CREST initiative, which has sought to promote the development of technologies that will be of benefit to ExoMars-like missions. This work of course may be complementary to wider research in this area.

2.2. Objectives

The objectives for this work were as follows:

- establish an initial scientific methodology for the automation of science assessment and planning based on terrestrial field practice
- prototype a system architecture that can support the concept of autonomous opportunistic science
- prototype elements of the methodology provided by the science team in order to establish the feasibility of this approach
- demonstrate the prototype system in a representative "Mars Yard" environment
- use the forthcoming ESA ExoMars mission as a target and source of operations and science requirements

Our primary task was to demonstrate opportunistic science in a representative ExoMars environment. A reference scenario was outlined in the early phases of the study, which provided a baseline for our work. The objective was to demonstrate that the mobile platform could traverse a rock field en route to a target destination and both detect and respond to targets of scientific interest that it encountered en route. The response was to occur at two levels: first to authorize close-up or high-resolution imaging of a target detected in a WAC image and second to plan and place a robotic arm if the target was sufficiently interesting. This required a number of key components including science assessment and response, replanning and resource monitoring, and robotic arm approach and placement. In addition we required a basic system to support this scenario, including ground-based planning and onboard software elements such as time line or plan execution.

In previous ESA-sponsored work an autonomous science assessment capability developed for a Martian aerobot prototype was used as the starting point for the autonomous science aspect of this work (Woods et al., 2008). However, the requirements for a rover-based assessment system are very different, so a new approach has been developed. The study team includes a planetary geologist, and a key feature of this work is our attempt to define a framework from which to build a hierarchical scoring system for scientific evaluation based on fundamental geological features. Although ExoMars is the initial target mission for this work, the framework approach is generic and could be used by any surface element. The intent is to base this model on terrestrial geological field practice (Stow, 2003) while considering the constraints associated with robotic exploration (Pullan et al., 2008). The framework is used in turn as a basis for our autonomous science assessment and response models.

In another ESA-sponsored activity, artificial intelligence (AI) planning and scheduling-based tactical replanning software called time line validation and control (TVCR) has been developed to support goal-based arbitration and time line replanning to support opportunistic science (Woods et al., 2006). Recent work at Aberystwyth University (AU) provided both the test environment or "Mars Yard" and a half-scale model of the ExoMars chassis E concept with a simple assembly representing a robotic arm.

3. TEST BED FRAMEWORK

TVCR is at the heart of the autonomous science concept and will be used to reason about the suitability of servicing science operations requests generated by the onboard science component. The robotic arm agent will provide the basis for an autonomous implementation for more detailed science assessment requests by supporting deployment of the robotic arm. The basic operations or usage model for the system is as follows:

- Nominal exploration time lines or plans are uplinked from the off-board mission planning element.
- The rover executes the planned sequence, which is mainly a traverse action between designated waypoints.
- At selected points the imagery collected during the traverse is assessed for science interest.
- If sufficient interest is detected, the science assessment and response agent (SARA) will request new data acquisition and time for itself to perform a more detailed analysis via the executive and TVCR.
- TVCR will assess the current plan, resource state, and mission priorities before recommending a go/no-go for the new opportunistic science request.
- The request may involve a close-up image activity or an actual arm placement on a target object such as a rock or outcrop.

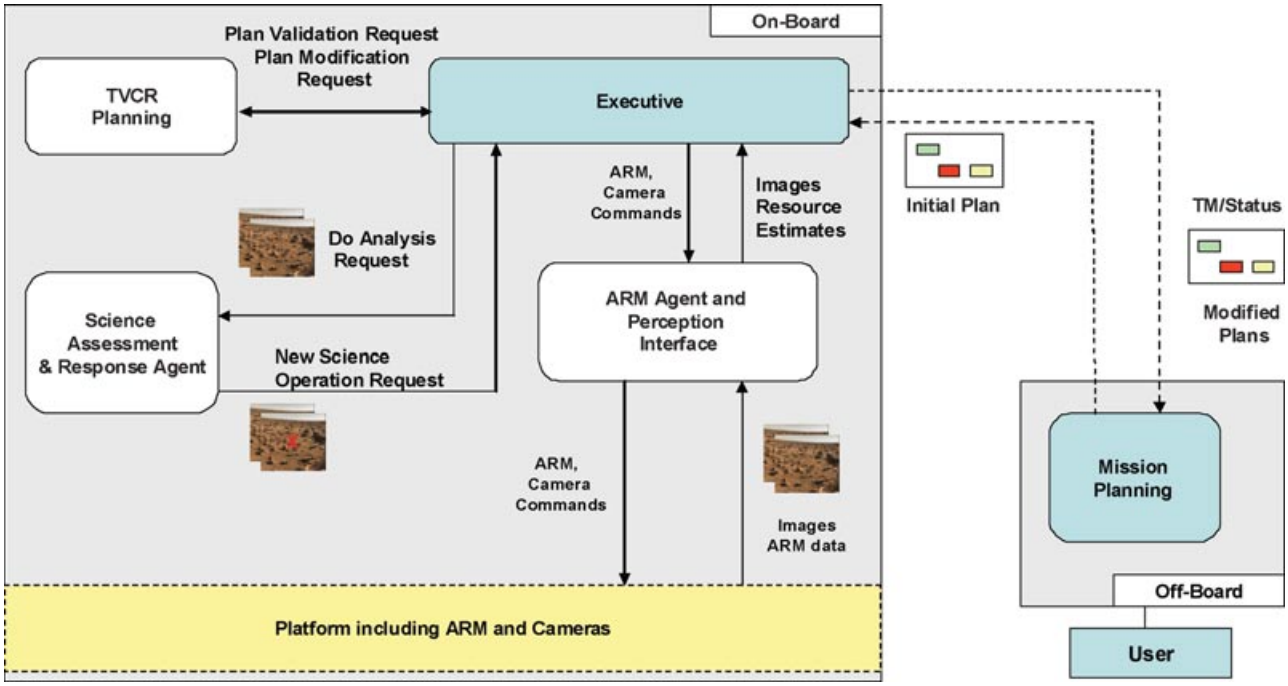

Figure 3. CREST system architecture.

Each of the primary components, namely science (SARA), planning (TVCR), and the arm agent (AAPI), is discussed in more detail in the following sections.

Figure 3 shows the architecture we have developed for the CREST demonstrator system, consisting of an autonomous science assessment component, closed-loop robotic arm approach and placement (arm agent and perception interface), and an onboard planner and scheduler TVCR.

4. SCIENCE ASSESSMENT AND RESPONSE

The SARA component is based on the underlying scientific scoring framework summarized below.

4.1. Science Assessment Framework

Geological features often appear complex and are influenced by a huge number of variables. In the field, human geologists mentally deconstruct what they see and draw on broader contextual input (the bigger picture) to help classify geological materials and the processes that act on them. Observations made in the field, aided by effective use of a hammer and a hand lens, provide an assessment of structure, texture, and composition, the basic ingredients for interpretation. This should also be the objective of a robot undertaking planetary geological fieldwork. In either case, interpretation relies on iteration because features seen from afar often look very different when viewed close up (sometimes unexpectedly so). This emphasizes both the importance of detailed close-up observations (payloads must be equipped with appropriate deployable instruments and tools for in situ work) and the need to incorporate reevaluation into the onboard autonomous routines.

For the purpose of this study, the following were considered necessary to establish a framework for a first-generation system based on autonomous science:

- definition of the fundamental attributes applicable to geological features
- mechanism for assigning scientific value to these attributes
- methodology for constraining scientific value due to external factors (i.e., context)
- requirements for rule-based algorithms
- facility for testing and evaluating the concept

As mentioned previously, the basic attributes used in field classification of geological features are *structure* (defined by parameters such as geometric shape, scale, orientation, and form), *texture* (parameters fabric, luster, relief, grain size, shape, and sorting), and *composition* (parameters color, albedo, specularity, and mineralogy). All these attribute parameters can be expressed as discrete constants or modeled as multivalued functions to reflect a variety of the real-world situations ("vitreous luster") and consequent uncertainty.

For this study, a feature database of commonly used parameters in terrestrial field geology was compiled for each basic attribute, with each assigned an arbitrary science value score (SVS) based on relative geological significance. Each target could be assessed and described using the elements from a set of possible parameters for each attribute. Once the appropriate parameters were identified for each target, their scores, defined in the database, could be used to build up a cumulative score for each attribute. At the attribute level, therefore, the accumulated scores represent indicators of "feature richness." The overall SVS of a science target is derived by combining all the features observed/measured, expressed as parameters for each attribute, and making adjustments based on data quality and context (knowledge of the geological environment in which the robot explores). At the start of the mission, a contextual model would be primed using prelanding data (regional geology from orbital mapping) and mission objective criteria (search for signatures of life). Following landing, the contextual model would be enhanced using ground truth from panoramic landing site surveys and in situ analytical measurements.

In general terms the *total* SVS of the target is a function of a number of derived parameters:

$$\mathrm{SVS} = f(As, At, Ac, Ax, Q, B),$$

where As is the overall structural attribute score, At is the overall textural attribute score, Ac is the overall compositional attribute score, Ax is the composite attribute score, Q is a quality factor, and B is a bias factor.

A simple method of calculating total SVS could be expressed as follows:

$$\mathrm{SVS} = \left(\sum \mathrm{A}s + \sum \mathrm{A}t + \sum \mathrm{A}c + \mathrm{A}x\right) \cdot Q \cdot B. \quad (1)$$

The composite score ($\mathrm{A}x$) is applied when certain, identified attribute parameters occur in combination. Note that the quality and bias factors Q and B are intended to enhance or diminish the overall score in much the same way a human geologist may apply these criteria in the field. Q will eventually be derived from the recognition algorithms and instrument parameters such as focus, resolution, and illumination. B will be derived from the contextual model and currently provides a simple means of weighting the significance of certain features with respect to the current environment.

Because this study was heavily constrained in both time and manpower for this topic, it was important to demonstrate a mechanism by which this could be achieved for a variety of situations using the most basic parameters and a simple scoring system. Consequently only *As* and *Ac* were implemented in SARA and restricted to only layering/no-layering and dark-toned/light-toned, respectively, at this time. Future versions of SARA will replace simple arithmetic weighting [Eq. (1)] with a more appropriate scoring system based on objectivist Bayesian probability as used by others (Pedersen, 2000).

Table I lists some examples of how one could derive a SVS for a target using the simple arithmetic method described. The interpretation assigned to each observation (feature classification derived from particular configurations of attribute parameters) is compared to table entries within a simple contextual model in order to assign the appropriate bias factor. How the observation maps to the interpretation will be the subject of further study into the general problem of modeling the impact of context.

Table I illustrates some interesting potential outcomes using this methodology. Low-level attribute parameters such as *structure.signature* and *texture.size*, if detected by the feature recognition algorithms, are assigned predefined scores as specified in programmable attribute lookup tables. These scores and the occurrence of certain *attribute.parameter* configurations (indicated in boldface in Table I) lead to feature classifications (column 1 in Table I) and the assignment of a composite score in some cases, which elevates the importance of the observation. Example 1 (cross-bedded sandstone) has a predictably high SVS of 1,245 due to its compound feature set and distinctiveness. Example 2 (salt deposit) has fewer features and would normally have an SVS of 210. In this case, however, a bias factor of 10 is currently assigned in the contextual model to detection

Table I. Examples of composite features.

Feature classification (derived)	Ax[a]	Attributes (As, At, Ac)[b] and attribute parameters	Quality	Bias	SVS[c]
Cross-bedded sandstone	1,000	structure.signature = "distinct" = 25 **structure.extent = "continuous" = 20** **structure.form = "planar" = 10** **structure.orientation = "multiple" = 50** texture.matrix = "clastic" = 20 texture.roundness = "subrounded" = 20 texture.size = "medium sand" = 10 texture.sorting = "well sorted" = 40 **composition.mineralogy = "quartz" = 50** Subtotal SVS = 245	1	1	1,245
Salt deposit	0	structure.signature = "none" = 0 texture.extent = "homogeneous" = 10 composition.albedo = "high" = 100 composition.color = "white" = 100 Subtotal SVS = 210	1	10	2,100
Carbonate	0	structure.signature = "none" = 0 texture.signature = "indistinct" = 5 composition.mineralogy = "dolomite" = 9,999 Subtotal SVS = 10,004	0.25	1	2,501
Vesicular basalt	0	texture.relief = "vesiculated" = 10 composition.signature = "distinct" = 50 composition.petrology = "basalt" = 5 Subtotal SVS = 65	1	1	65

[a]Composite score Ax due to combination of attributes highlighted in bold.
[b]Parameters and assignments derived from the feature database (i.e., As, At and Ac data tables).
[c]Final (modified) SVS [see Eq. (1)].

of "salt deposits" as these have yet to be found but are expected at the landing site. Thus the salt deposit scores higher than the cross-bedded sandstone. Example 3 (carbonate) is not expected at the landing site so has a bias factor of one. However, the detection of carbonate overrides the previous two examples even though the data quality is poor. This is due entirely to the high SVS assigned to the carbonate composition (i.e., 9,999). Example 4 (vesicular basalt) scores lower than all previous examples even with an unambiguous petrological interpretation. This is due to a combination of uniqueness and feature richness of the previous examples.

4.2. Science Agent

The concepts outlined above have been used to create the basic SARA architecture outlined in Figure 4. The structure of the architecture is built around three levels of processing: candidate target area extraction, geological attribute analysis, and finally the fusion of results factoring in geological context. Our analysis was closely tied to the scale at which the image data were retrieved.

For geological fieldwork, activities can be classified on the basis of three practical working distances between observer and target: proximal (~100 cm), macroscopic (~10 cm), and microscopic (~1 cm). This range can be considered to be synonymous with the immediate radial "working zone" of a stationary human field geologist who is suitably equipped with tools of the trade (hammer, field lens, portable analyzers, samplers, etc.), and these terms are adopted to distinguish from activities beyond the physical reach of the observer (i.e., "remote" sensing) that require mobility to reach. In the planetary context, the

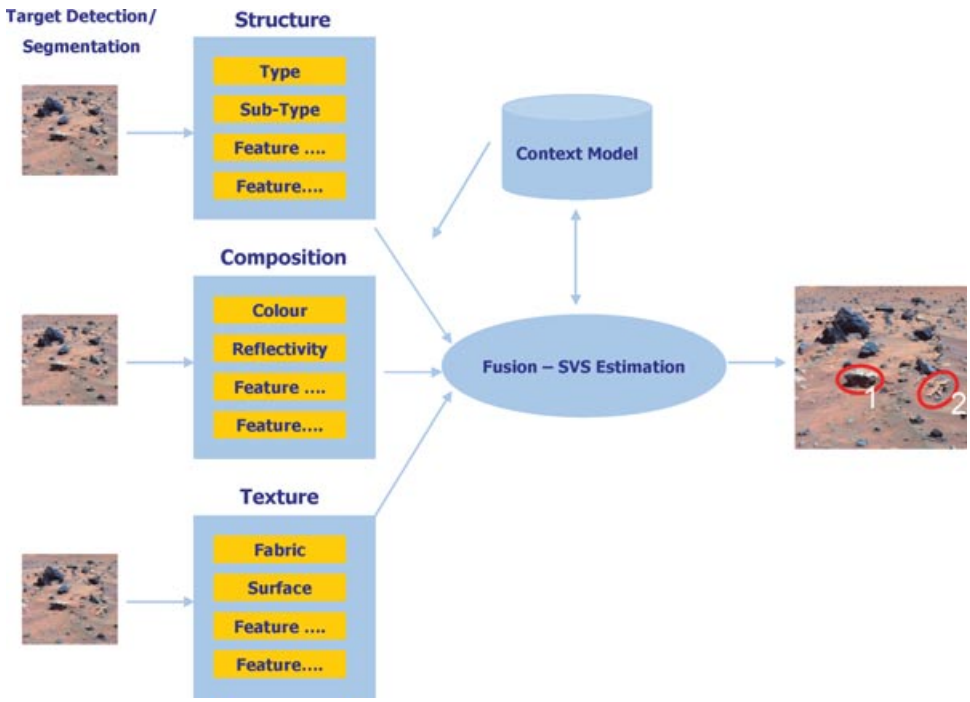

Figure 4. SARA architecture: Target assessment is built up through the analysis of individual attributes and the parameters that define them. Context information is used to further qualify the final target score.

analogy applies equally to static planetary landers or stationed mobile vehicles equipped with robotically deployed scientific payloads.

This study considered only proximal and macroscopic working distances. How the images are actually analyzed depends on the scale at which the input image was taken. For example, initial images taken at the proximal distance provide macroscopic detail (standoff image scale) and are assessed for target rock features on the basis of general morphology, whereas a more detailed analysis of grain and bedding macroscopic features is applied to high-resolution images (close-up image scale). The same basic process flow applies in each mode, i.e., segmentation, analysis, and fusion, but the algorithms used in the analysis in particular differ. This approach supports the hierarchical analysis characteristics of field practice, in which analysis starts off light and grows in rigor and usually narrows in focus in a series of discrete stages.

In this instantiation we focused on developing algorithms that could detect a variety of bedding types and basic morphological analysis in order to assess the target structure. For composition we considered albedo and reflectivity, and for the texture we sought to detect granules. We restricted use of the architecture to "standoff" and "close-up" image scales. Standoff is represented by the immediate, proximal, terrain view available in the WAC images. In effect this captures all objects 1 m from the front section of the rover at medium resolution. Close-up is a more detailed view of a macroscopic rock or outcrop features provided by a high-resolution close-up image.

Given the timescale of the work, we did not model context in a sophisticated way, so the final SVS for each candidate target was simply a weighted fusion of a limited number of individual attribute scores. Learning techniques were developed but testing was limited to processing stand-alone image groups from MER images as time constraints prevented evaluation in the final field trials.

In the standoff phase the input images are preprocessed in order to detect likely target candidates, i.e., rocks or outcrops. To achieve this we relied on a number of basic image processing techniques and applied them in the following way: initially the image was smoothed using a morphological [Eq. (2)] and Gaussian smoothing algorithm; the results were

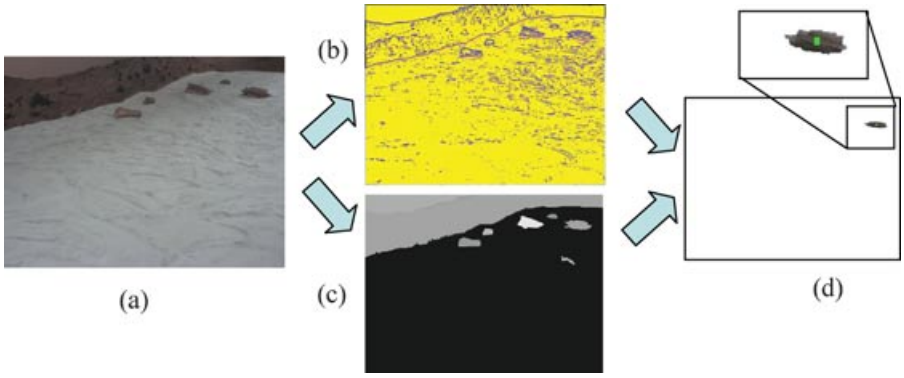

Figure 5. Standoff image processing sequence: (a) original WAC image, (b) feature extraction image, (c) result from smoothing and segmentation, and (d) result from the combination of (b) and (c) and the detection of an interesting target (the segmented area around the target is reduced to remove boundary artifacts). During standoff analysis the contextual model is used to denote "what" is interesting using a reduced SVS parameter set. The feature vector shown in (b) is compared with expected values and assessed for terrain "roughness" characteristics. During the trials this was set at a roughness value of 0.3 relating to the rock ridge and channel feature density (potential bedding) and albedo range. If the predetermined thresholds are breeched, a positive request for further analysis will be issued. Note that this thresholding of the derived feature vectors does not use the SVS function derived in Eq. (1), which is used only in level 2 close-up analysis in this instantiation of the system.

analyzed using a graph-based, region-growing algorithm (Felzenszwalb & Huttenlocher, 2004), and by applying a threshold it was possible to identify the rocks from the soil. The original color image was then gray-scaled ready for feature classification (Shaw & Barnes, 2003) using a double differential method [Eq. (3)]. Targets of interest were identified by combining the outputs from both processes, the segmented regions being used as a mask on the feature image. This method of analysis and identification proved adequate for our target environment and the subsequent integration trials (see Figure 5) and also images of Mars obtained from previous missions (see Figures 20–23 later in this paper). Robust rock and outcrop detection under a variety of conditions is of course not trivial and has been the subject of work by other groups (Thompson & Castano, 2007).

Once regions of interest have been defined, the next step is to carry out an attribute analysis:

$$I^F = I^O_{(i,j)} \bullet m = \left(\left\{\left[\left(I^O_{(i,j)} \ominus m\right) \oplus m\right] \oplus m\right\} \ominus m\right). \quad (2)$$

Equation (2) is the calculation for the morphological smoothing. The mask (m) is applied to each pixel in the original image performing and opening followed by a closing operation (Gonzalez & Woods, 1992). The opening operation removes small light details, leaving the overall gray levels relatively unchanged, and closing removes dark details from the image, again leaving the overall gray level the same.

$$I^F_{(i,j)} = f\left(\frac{d^2z}{dx^2}, \frac{d^2z}{dy^2}\right) + f\left(\frac{d^2z}{dxy}, \frac{d^2z}{dyx}\right). \quad (3)$$

Equation (3) is the calculation of the feature density, using the radius of curvature at a point that is calculated as a function of the double differential along the x, y, xy, and yx axes. The value and sign of the components denote the direction of a curve allowing feature determination.

Figure 4 indicates that attribute analysis is dependent on the detection and classification of primary features such as color, texture, and structure. The detection and classification of these features is complex and will require a long-term effort to build up an extensive capability.

This study concentrated on the detection and analysis of bedding and to some extent granularity and albedo in the close-up analysis. This process also started with morphological and Gaussian smoothing, followed by a combined Sobel edge detection and morphological gradient detection [Eq. (4)]. The detected lines were binarized and thinned before being analyzed for direction and length. The profile of the direction and length of the lines was then examined in order to determine whether any bedding was

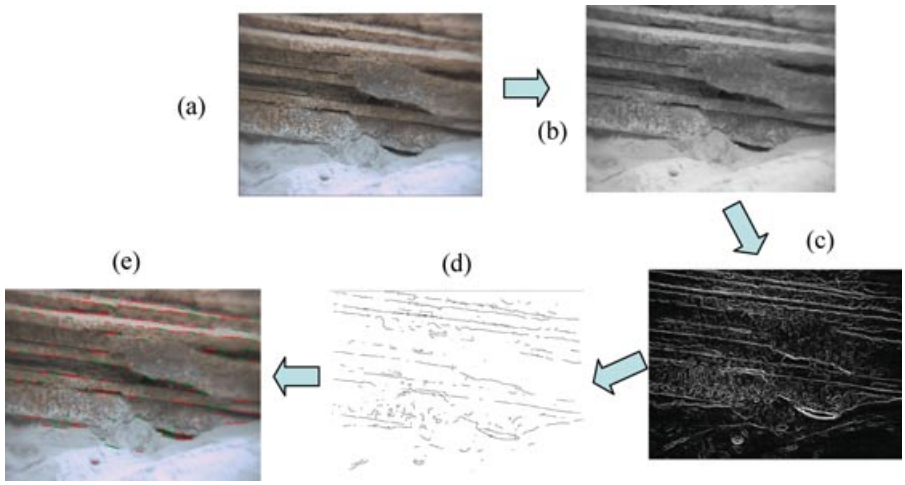

Figure 6. Close-up images processing sequence: (a) original zoom image, (b) gray-scaled and smoothed image, (c) Sobel edge detection and morphological gradient combination, (d) line filtering and thinning, and (e) line direction and length classification superimposed onto the original image.

present and also its type and direction (see Figure 6).

$$I^F = I^O_{(i,j)} \bullet m = \left(I^O_{i,j} \oplus m\right) = \left(I^O_{i,j} \ominus m\right). \qquad (4)$$

Equation (4) is the calculation for the morphological gradient. The mask (m) is applied to each pixel in the original image performing a dilation and erosion on the image pixels (Gonzalez & Woods, 1992).

A similar technique is used to search for granules, in which the image is processed with a Gaussian smoothing and morphological gradient technique before an incremental threshold procedure is applied. The result is then analyzed for granules of various sizes; the profile of their size and distributions is interpreted to give a result.

As the color cameras had been calibrated only for distortion and not for color correction and reflectance, only an approximation of the albedo values of the rocks could be generated; this was achieved by gray scaling the original image and taking the average gray intensity value for the rock. As the actual lighting conditions varied between the tests, the albedo was measured but the value was used only to identify objects that were "unique."

All the methods implemented could be configured through a parameter script to allow the adaption of the analysis and emphasis of the object attributes. Clearly, the attribute classification will be improved if more feature identification methods are included in the processing repertoire. Figures 5 and 6 show the intermediate results from processing at both standoff and close-up positions.

5. REPLANNING AGENT

Classically, in AI planning, replanning is defined as the task of taking an existing plan, together with new goals, and constructing a new plan that achieves both the original and new goals. This problem is as hard, in general, as planning from scratch to achieve all the goals (Nebel & Koehler, 1995). In the context of the autonomous scientist, the replanning problem differs from this classical definition in several ways. First, a characteristic of the task-oriented behaviors that drive the science missions is that it is natural to describe goals not in terms of the states that must be achieved but in terms of the actions that must be executed. This is because the science-gathering actions have only one purpose, which is to acquire the science that is the true objective of the mission, so it is natural to specify the goals of a scenario by specifying the actions that must be executed. Second, there are typically few choices in the actions from which the plans are to be constructed, but more choices about the ordering of the critical science-gathering actions. This makes the problem more like a scheduling problem than a classical planning problem. However, the order in which science-gathering actions are executed determines precisely which supporting actions must be executed, and in what order, to link these activities

into a coherent sequence, and it is these actions that give the problem a planning character. Third, the most important constraint on the achievement of a successful replan is not the achievement of the goals, but the management of the limited resources.

These three factors contribute to shaping the replanning strategy in TVCR and in making a replanning strategy more efficient than planning from scratch. In the autonomous scientist's work, only the replanning functionality of TVCR is of significant interest: the other functions, including monitoring the execution of the time line, remain active but are not central to the focus of the research (Woods et al., 2006). To perform its replanning, TVCR reasons with a model of the activities that are available in the domain. This model is an action-centric model, providing declarative descriptions of the actions that the executive can be called on to perform. The model is written in planning domain description language (PDDL) (Fox & Long, 2003). The benefits of such a model are that it is relatively easy to extend with new actions, or to modify the descriptions of existing actions, in a form that closely corresponds to intuitions about the behaviors. Furthermore, PDDL is a widely used language, so its use offers access to a wide selection of existing planning and plan-manipulation tools. PDDL allows actions to be described in terms of their preconditions and effects. The action models used for this application include the primitives for traverse, approach target and for each of the basic science operations that can be performed.

The most important functionality on which TVCR depends is plan validation (Fox, Howey, & Long, 2005). This allows TVCR to take a time line description, either supplied from the ground or partly constructed onboard, and validate it. The validation process identifies predicted flaws in the execution, indicating where in the time line they occur and what their cause is.

In the context of the current work, a request to replan arrives at TVCR in the form of a fragment of a plan (a connected sequence of actions, but typically with unsupported preconditions) that is to be inserted into the time line. This is because the replanning request is triggered by the recognition, in the science agent, that there is a previously unrecognized opportunity to gather scientific data. The fragment has an associated priority value reflecting its estimated relative science value in comparison with the fragments that form the current time line. TVCR handles the request by inserting it into a plausible slot in the time line, identified by selecting the point at which the rover is located most closely to the site at which the fragment is to be executed, following other activities at the same site. Usually, this attempted insertion will generate plan flaws due to interactions between the newly inserted actions and the actions that were already on the time line.

TVCR has a series of strategies available to it to handle the flaws that can arise. These include removing low-priority actions when there is a shortage of resources such as power or data storage, delaying actions when there is a conflict between demands on fixed resources (such as instruments), and adding in support actions to ensure the coherence of the execution trace of the time line. TVCR applies these strategies in a fixed order. The resolution of some flaws can create new ones, and these are handled alongside the others. If the process of resolution fails to generate a valid time line within a fixed number of iterations, TVCR falls back on a fail-safe strategy of stripping down the original time line until the activities that remain fit within the resource envelope available and represent an executable plan. In this way, TVCR can ensure that there is always a valid time line awaiting execution (although it might be empty) and can add opportunistic plan fragments to the time line as they arise, while respecting the constraints on resource use and on correct execution of the plan. TVCR is deliberately designed not to attempt to perform full planning. Its behavior is precisely limited by the priorities that govern the order in which fragments may be included in or excluded from the plan. Further, it is restricted in the degree of search it may perform to find a repair. These constraints allow TVCR to operate within tight operational bounds on its performance: it does not build a large search space to explore alternative choices. They also ensure that TVCR respects decisions made by an external human agent about the order in which activities should be considered. This gives us significant power over the choices TVCR will consider, without being burdened with the need to model all of the constraints that lead to the development and ordering of those choices.

The use of plan fragments as an organizational structure for activities that are linked in the construction of plans makes the models TVCR uses similar to a single-layer hierarchical task network (HTN) (Erol, Nau, & Hendler, 1994). The expansion of a fragment replaces it with the body, which is always a uniquely specified collection of actions in a partial order. In contrast to HTNs, each fragment has only one

expansion. The expansion allows actions to be incorporated into the plan, together with constraints on their relative positions. Another contrast to HTN planning is that the plan can be modified by the addition of individual primitive actions (or even sequences of actions) that act as bridges between the plan fragments, ensuring that the state of the system is correctly prepared, ready for the execution of the next plan fragment.

The example in Figure 7 is a very simple illustration of the process of plan repair. In this case, the original plan consists of two fragments, each containing a traverse, a panoramic camera imaging action, and a standoff or level 1 science analysis action. If, following the first level 1 analysis, the science agent determines that an opportunity is available to perform a new science data gathering task, the new fragment shown in Figure 7 will be created. TVCR is then invoked in order to attempt to insert the fragment into the plan. TVCR recognizes that the precondition of this fragment includes the need to be at the appropriate science target site and will insert a minor traverse action (in stand-alone TVCR tests only not in final trials) to adjust the position to meet that precondition. It identifies that the best place to insert the opportunity is immediately, because this is the point at which it is closest to the target. The remainder of the plan is moved to make time for the opportunity to be inserted.

Several issues might affect this process. First, if the rover is committed to other activities in the same location, then TVCR might choose to insert the opportunity after one or more of these activities. Second, TVCR considers the overall constraints for the plan, in terms of time, memory, and energy. If the insertion of the opportunity would cause the new plan to exceed any of the limits on these resources, then TVCR will remove fragments from the modified plan, starting with the lowest priority fragment that has not yet been executed. Typically, opportunities will be given lower priority than the original plan fragments, but this need not be the case. Thus, if the second fragment were lower priority than the opportunity, TVCR could remove it in order to free sufficient resources to allow the opportunity to execute. TVCR will take into account the additional resource requirements of the supporting activities, such as the additional minor traverse in this case.

Planning and plan modification acts at a different level to the function of control languages such as TDL (Simmons & Apfelbaum, 1998) and RMPL (Ingham, Ragno, & Williams, 2001), but these levels interact closely. Control languages allow flexible programs to be specified for robotic platforms, with the intention that an executive should be able to interpret the commands of the language at execution time in order to manage the uncertainty inherent in the interaction between a robot and its environment. However, the programs written in these languages are the result of careful planning of the activities that the robot is intended to execute. Thus, planning can be seen as an analog of the programming activity that precedes the execution of programs written in these languages. In our implementation, the functions of a TDL or RMPL control program are subsumed by a purpose-built executive implementation that performs the necessary control. This component is not designed to be a fully robust solution so does not contain the full functionality of execution monitoring or the flexible execution of the form supported by TDL and RMPL.

Several other researchers have explored ways to provide a planning function in the face of the need for plan modification or plan failure. Among the most relevant work is that of Chien and his team (Chien, Smith, Rabideau, Muscettola, & Rajan, 1998; Rabideau, Knight, Chien, Fukunaga, & Govindjee, 1999), using ASPEN and its reduced form, CASPER. Chien and his group have explored the use of onboard planning technology to select and direct observation sequences for EO-1 (Sherwood et al., 2005). Their approach involves the use of a carefully constructed and complex onboard model that can be used to plan new observations as the goals arise.

The contingency planning work explored by Bresina, Golden, Smith, and Washington (1999) using CRL (contingent rover language) is also very relevant. Using CRL, the authors construct plans that include alternative paths triggered by circumstances that might arise during execution. These are very similar to the opportunity fragments used in TVCR. The plans they construct place the contingencies at specific points during execution (identified by particular circumstances observed via sensors). This is a slightly different view of the plan structure and requires that contingencies be considered for execution without exploring the impact on later resource availability. This means that the local modifications to the plan cannot compete with resource allocations to later parts of the planned activity. To allow later plan elements to be dropped in favor of earlier opportunities, it would be necessary to make the later elements

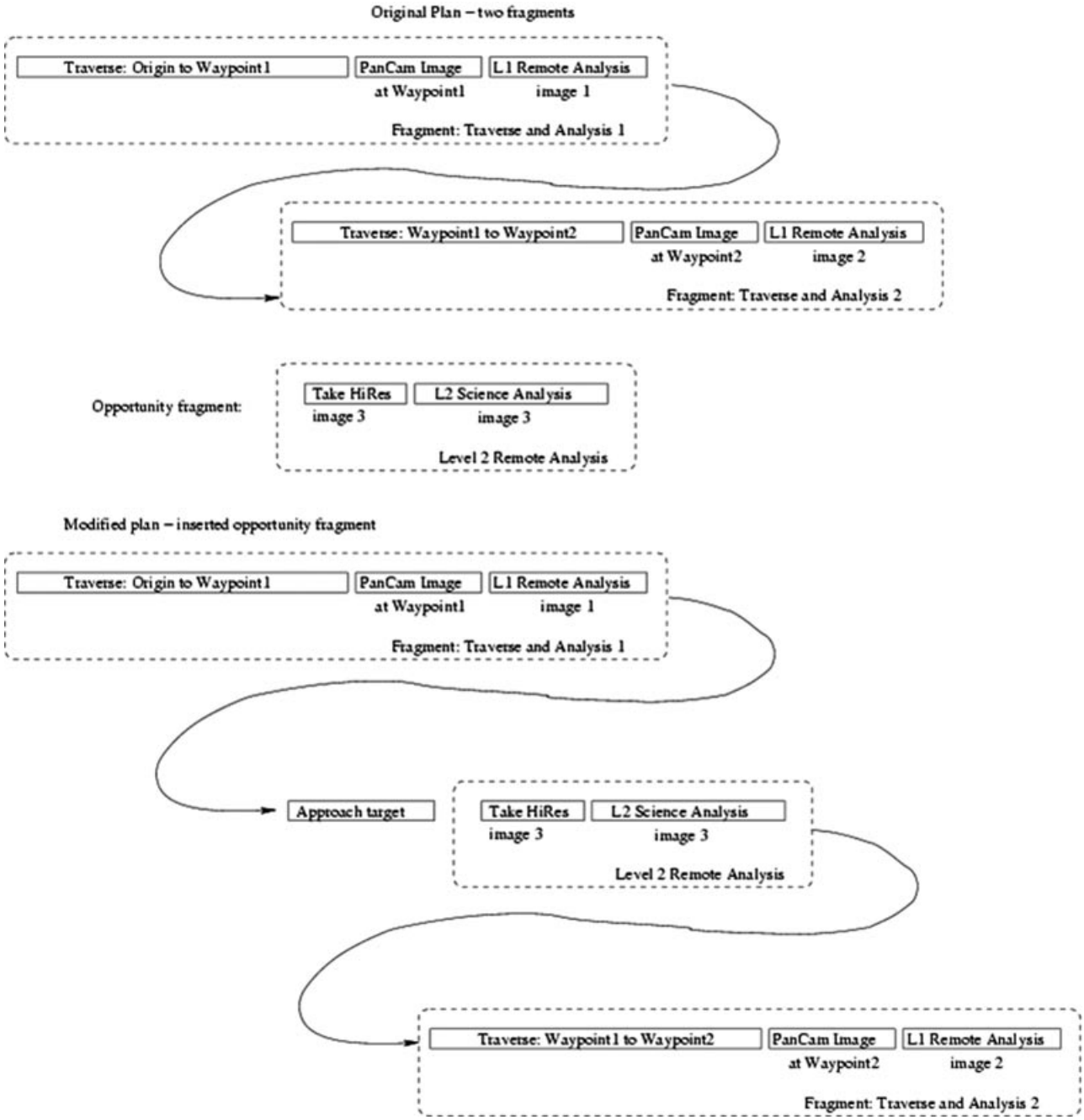

Figure 7. Simple plan modification by TVCR. The initial plan contains two fragments, including six actions (top of figure). There is a single opportunity fragment to be inserted. This is placed between the two fragments, and an additional action is inserted to link the fragments into a coherent plan.

contingent on there being sufficient resources to complete them. The most difficult aspect of using this mechanism to control the opportunity insertion is to anticipate the ways that current execution decisions could impact planned future operations without having an explicit review of those future operations. The CRL mechanisms allow current execution decisions to be affected by current state but do not attempt to review the future planned operations or to modify them ahead of execution. TVCR considers the entire planned sequence of operations when attempting to insert new opportunistic fragments and makes the decisions based on the anticipated requirements for resources of those operations.

6. ROBOTIC ARM APPROACH AND PLACEMENT PLATFORM

Key hardware components within the autonomous arm agent and perception interface (AAPI) of the CREST architecture are the panoramic cameras and zoom high-resolution camera (pancam), the pancam pan and tilt mechanism, the robotic arm, and the locomotion chassis with associated onboard computer and electronic interfaces (Figure 8).

Upon instruction from the executive, a stereo image pair is captured using the pancam WACs. Our demonstration scenario required an overlapping sequence of image pairs to be captured via the autonomous operation of the rover's pancam pan and tilt mechanism. SARA then examines one image (typically the left-hand image) from each image pair.

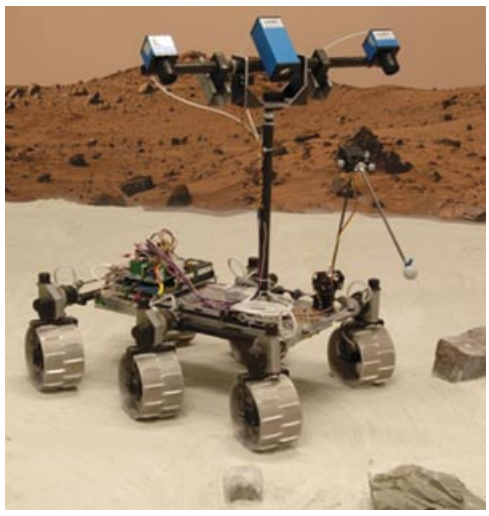

Figure 8. A close-up of our demonstration rover platform.

Upon identification of a science target, the image pixel coordinates of this object (e.g., the rock's centroid) are communicated to the AAPI. SARA can request a zoom image of the rock in question to confirm a science target hypothesis. The AAPI accomplishes this by calculating a correction to the pancam pan and tilt mechanism orientation so as to center the science target in the zoom camera's field of view. A rock zoom image can then be captured. Using the science target WAC-captured image pair and stereo triangulation, the 3D position of a science target is calculated relative to the rover. This allows the pan and tilt orientation to be calculated for the zoom image capture activity and allows a science acquisition "cost" to be calculated. This is based on the power and time that would be required for the rover to perform placement onto the science target location.

The cost information is used by TVCR to assess the resource implications of an opportunistic science activity. If a "go" is given, a rover traverse can occur to place the science target within arm's reach, although this was not incorporated in the final field trials. The AAPI requests a WAC image pair and uses its stereo triangulation and arm kinematics to confirm target reachability and calculates an appropriate arm configuration for instrument placement.

A safe instrument placement trajectory and contact region on the science target can also be determined by generating a mini-DEM (digital elevation model) of the science target. The autonomous AAPI has required a number of algorithms to be designed and implemented, including camera image distortion correction and rectification, disparity map generation, stereo triangulation, pan and tilt mechanism pointing, science target acquisition cost calculation, arm reachability, and safe instrument placement trajectory determination and execution. Our current research is focused upon an eventual field programmable gate array (FPGA)-based implementation of the AAPI functionality.

6.1. AAPI Details

The AAPI algorithm contains the following steps:

- capture pancam stereo images, remove lens distortion, and rectify
- obtain pancam (2D) image pixel coordinates for candidate science targets

- generate disparity map, and use stereo triangulation to create science target DEM
- verify that targets' coordinates are within robot arm reachability envelope [YES↓,NO↑]
- identify (optimal) instrument placement regions within reachable DEM region
- identify correspondence between placement regions and science targets [YES↓,NO↑]
- generate science targets' surface normals and placement trajectories
- check for collisions along placement trajectories [YES↓,NO↑]
- estimate costs (time and power) to place instrument on science target [YES↓,NO↑]
- if all OK, then "get sample"—wait for next candidate science targets

At key points within the algorithm, critical decisions have to be made that will result in a continuation to the next step (denoted YES↓) or the algorithm will detect a problem with the candidate science target (e.g., unreachable with the arm). Such a decision (denoted NO↑) will cause the AAPI to flag that a new science target is required.

The AAPI pancam was calibrated prior to conducting the field trials. A standard camera calibration method (Bouguet, 2008; Zhang, 2000) was employed that required a 16 × 16 checkerboard target to be imaged by both the left and right WACs. These data allowed both the extrinsic and intrinsic properties of the cameras and their optics to be obtained. The pancam calibration data allowed the lens distortions to be removed after image capture and prior to stereo image processing.

The AAPI stereo triangulation algorithm required a simplified epipolar geometry to be observed, and hence any captured WAC images had to be rectified (see Figure 9). Rather than implement a "yet-another-disparity" algorithm, we wished to investigate the performance of a state-of-the-art approach that showed good performance when compared to other algorithms and was able to deal with occlusion problems (a situation that is quite probable in a Martian "rock garden"). We based our disparity map generation on the cooperative algorithm for stereo matching and occlusion detection by Zitnick and Kanade (2000). We found that this algorithm performed well provided that good (close to solution) initialization *minimum* and *maximum* pixel disparity values were selected a priori. The major problem with such an algorithm is the large computation time. Although this may not be an issue for terrestrial applications, when using disparity algorithms onboard an autonomous rover it must be noted that processing memory and power are very limited (for ExoMars typically of the order of 10s of megabytes of memory and a 100-MHz clock rate). For our experiments we found that simply creating an adjustable crop window around the given science target pixel coordinate considerably reduced the overall computation time (typically 40 s to generate a disparity map using a 1,024 × 50 pixel crop window for both left and right WAC images, running on a 1.2-GHz Intel Pentium M processor with 0.99 GB of RAM as used in the experiment).

Once a disparity map had been generated, then a simple process of stereo triangulation was performed using the obtained science target left_image$_{x,y}$ and

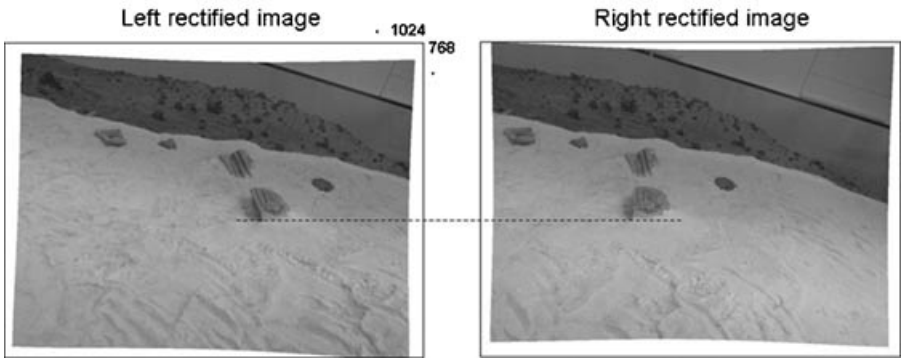

Figure 9. Example of left and right WAC image rectification. The black dotted line shows an example of the resultant simplified epipolar geometry.

right_image$_{x,y}$ pixel values and the previously obtained WAC extrinsic parameters.

Having obtained the 3D position of the candidate science target relative to the left WAC origin,[1] this position had to be transformed from the camera coordinate origin to the origin used by the rover arm kinematics (in our case, at the base of the arm).

Prior to conducting our field trials, both the rover pan/tilt mechanism and the rover arm were calibrated. This involved obtaining measurements of the *potentiometer-joint-offset* (PJO) for each servo mechanism used in the pan/tilt and arm devices This offset relates to the difference between the geometric 0-deg angular position of a given rotation joint axis and the actual 0-deg angular position as measured by the joint's servo mechanism potentiometer. This type of joint offset can be introduced during the device assembly process, for example, when the arm is typically constrained in a rig to align the joints and when the measurement encoders (potentiometers) are subsequently mounted to each joint.

Calibration of the *kinematics-joint-offset* (KJO) for each arm revolute joint was undertaken. The KJO for each joint represents the angular difference between the theoretical (kinematics) joint value and the actual joint value required to place the real arm's end point at a given target position in 3D space. The major contributor to this angular difference is arm deflection (due to link bending and torsion) when subjected to gravity. KJO calibration involved specifying a set of calibration 3D points (relative to the arm origin) within a given theoretical arm reachability region. These points were converted to joint angles using our arm inverse kinematics model. The arm was then commanded to execute these joint angles, and thus the arm was moved to the calibration 3D points within the real world. For each joint set (*base-angle*$_i$, *shoulder_angle*$_i$, *elbow_angle*$_i$), the actual 3D position of the arm's end point was measured using an external measurement system. A Vicon motion capture system was employed for these measurements, and this captured in real time (120 Hz) the Cartesian position (accuracy $\approx \pm 0.1$ mm) of a set of passive reflective spherical (25-mm-diameter) markers placed on the rover arm. The captured data were then used as inputs to our KJO calibration algorithm, which was based on a model of the arm's forward kinematics for each calibration 3D point and a nonlinear Levenberg–Marquardt least-squares error minimization approach. The calibration algorithm calculated a single value for each KJO, which were found to be base joint offset $= -1.185$ deg, shoulder joint offset $=$ 3.045 deg, and elbow joint offset $= 5.084$ deg. These KJO values were then incorporated into our arm's forward and inverse kinematics models for subsequent use by the AAPI.

A further calibration that was performed was to obtain the (left) camera origin to arm origin geometric transformation, so that the 3D position of the candidate science target relative to the arm could be obtained. An external measurement approach could have been used to accomplish this calibration, but we wanted to investigate the application of a potential in situ (i.e., on Mars) calibration method. Here the pan and tilt device was oriented so that images of the arm's end point could be captured in both the left and right WAC. This pan/tilt orientation was referred to as the *zero pan/tilt state* (the values used were pan $=$ 180 deg, cameras looking forward; tilt $= 45$ deg, cameras looking downward relative to the horizontal). When in this orientation the arm was moved to six different end-effector locations, where each location would be within the typical arm/rock contact operating envelope. A fiducial marker had been placed at the end of the arm, and it was important to be able to view this marker in all six arm positions and in both the left and right WAC images. For each arm location, our inverse kinematics algorithm was used to obtain the 3D position of the fiducial marker relative to the arm's Cartesian origin.

Each of the six image pairs (left and right WAC) of the fiducial marker was then rectified, and using the stereo triangulation algorithm, the 3D positions of the fiducial marker were obtained relative to the left camera Cartesian origin. A least-squares Helmert transform–based algorithm was then applied to the fiducial 3D camera and 3D arm data to obtain the camera origin to arm origin geometric transformation parameters. A six-degree-of-freedom (DoF) transform is created (plus scaling) that can be used to map a candidate science target from camera 3D space to arm 3D space. No additional camera origin to arm base origin transform parameters are required.

Figure 10 shows an example image pair used as part of this camera to arm origin transform work. The resultant transform was used as part of the AAPI processing during the field trials, and it was shown

[1]It should be noted that by using additional pixel values surrounding the original science target pixel value, a set of 3D positions can be calculated to yield a small science target rock DEM.

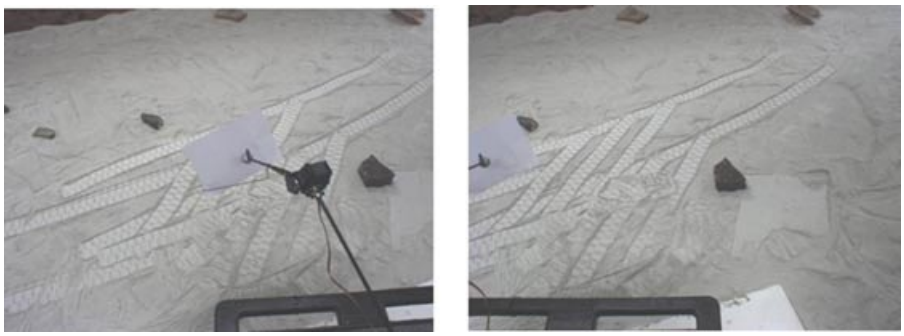

Figure 10. Left and right WAC images (prior to rectification) showing the fiducial marker attached to the arm end point. A piece of paper was placed behind the marker to aide observation.

that such an approach could be used for in situ camera/arm calibration should it be required (the in situ–captured stereo camera images could be processed on Earth, and the resultant transform uploaded to the rover on Mars).

Before the generated Helmert transform could be applied during the field trials, any calculated candidate science target in 3D camera space had first to be transformed to the zero pan/tilt state, and another trigonometric-based transform was created for this purpose.

Figure 11(a) shows an example of applying the zero pan/tilt state transform. This shows a plan view, and the science target in camera space is shown by the crosshair. "RP" shows the (relative) orientation (pan) of the left/right WAC baseline separation when observing the science target. "TP" is the result of transforming the WAC baseline separation to the zero pan/tilt state. The resultant transformed target x and z Cartesian values are shown with the dashed lines.

Figure 11(b) shows the same example but now in side view. Again the science target in camera space is shown by the crosshair. "RT" shows the orientation (tilt) of the WACs relative to the science target. "TT" is the result of transforming the WAC tilt to the zero

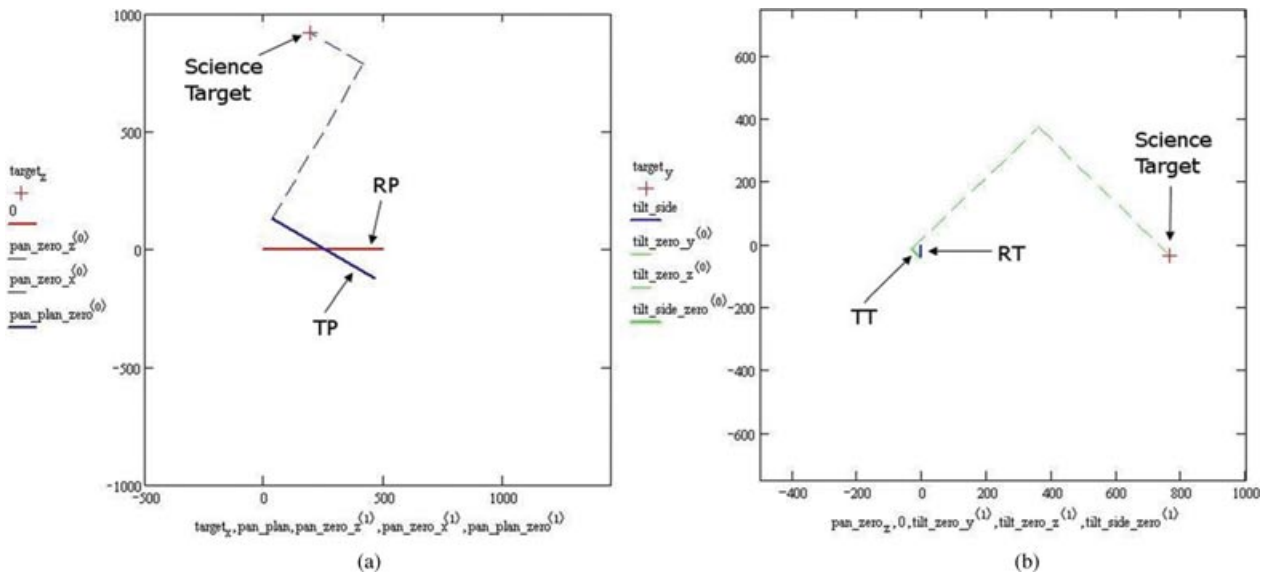

Figure 11. Zero pan/tilt state transform example: (a) plan view and (b) side view. Axis units are in millimeters. Coordinate system origin here is the camera origin for the left WAC.

pan/tilt state. The resultant transformed target y and z Cartesian values are shown with the dashed lines.

The AAPI incorporated additional algorithms for zoom camera pointing and (as mentioned previously) standard trigonometric-based forward and inverse kinematics algorithms for arm end-effector 3D position determination and end-effector positioning, respectively. The zoom camera pointing algorithm was used so that when a candidate science target had been found within a left WAC image, new pan and tilt joint rotation values could be calculated in order to slew the cameras to an orientation such that the science target was now along the principal axis of the zoom camera. Prior to moving the arm end point to a candidate science target position, the arm kinematics algorithm (together with the generated rock 3D data) was used to check that the target was within the reachability envelope of the arm. The generated rock DEM can be processed for (optimal) placement regions; in our case the simplest approach we have experimented with involves searching for the largest (relative) planar region (with respect to the horizontal) and then finding the centroid of this region using standard image processing methods (see Figure 12). Our AAPI approach does not require rock boundaries to be determined as the largest planar region detection process effectively segments one placement region from another.

It must be emphasized that SARA provides science target data (i.e., the left WAC pixel coordinates for the centroid of the identified rock science target), based entirely on the rock's science potential and not on the practicalities of actually obtaining a science sample from the rock in question. Naturally occurring rocks and an instrument head (end effector) typically have complex geometries, and it may not be possible to deploy multiple instruments to the same SARA-identified rock centroid location without rock/instrument head collision. The ability to deploy multiple instruments to the same point on a science target rock is regarded as an important operations capability. However, the computational overheads in ensuring that for every individual instrument placement and rock contact, there would be no collision between the remainder of the instrument head and any other part(s) of the target rock would be undesirable.

Figure 12. Plan view of "rock garden" DEM data that have been processed to obtain the rock planar regions for arm placement purposes. The black crosshair denotes the centroid on the largest contiguous planar region.

The AAPI approach is to identify "optimal" placement regions on or as near as possible (based on a simple Euclidean metric) to the original SARA-identified science target. Hence, for example, a placement target location that is surrounded by a flat (planar) region would be preferable to a placement target location that is at the bottom of some pit-type ("concave") surrounding region. The AAPI generates placement target data based on geomorphological analysis methods developed for geographical information system (GIS) technology (Wood, 1996). Given a science target rock DEM and a start DEM grid cell (derived from the SARA science target data), a local window (usually 3×3) is passed over the DEM. The central DEM grid cell value is compared to its neighbors, and based on the relative DEM height values a number of different 3D rock surface features can be determined. Figure 13 shows examples of six 3D rock surface features. In addition to identifying rock 3D surface features relative to the horizontal, the AAPI can also vary the feature gradient so that, for example, a planar region at an angle of 45 deg relative to the horizontal can be identified.

Typical arm operations involved executing a gross arm motion to a safe "standoff" position using joint-by-joint control (i.e., one joint at a time was moved). A simple straight-line trajectory from this safe position to the target rock contact position was then computed based on the work of Taylor (1979). Final arm motion along the target trajectory was accomplished using joint-interpolated motion (i.e., all joints were moved simultaneously, and each joint motion velocity was zero when the target position was reached).

Based on a priori information obtained regarding motor speeds and power consumption for the rover

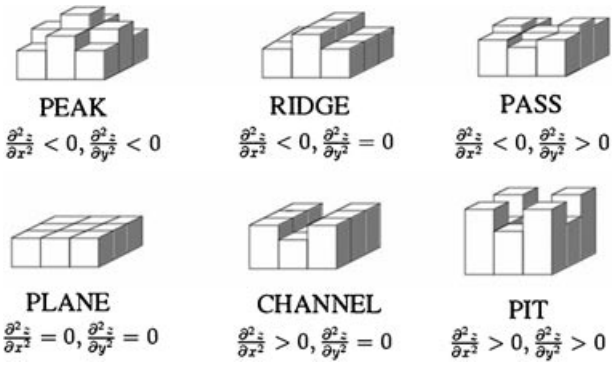

Figure 13. Example 3D rock surface features and their associated DEM height derivative relationships that can be used to determine an optimal instrument placement region given a SARA-identified science target.

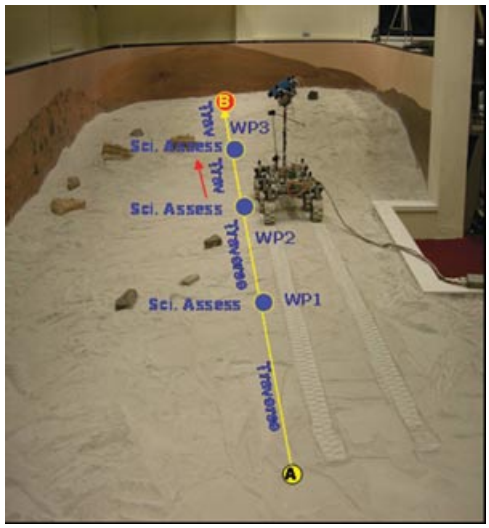

Figure 14. Demonstration scenario showing nominal route (waypoints WP1–3) with arrow pointing to the rock of most interest at WP3.

chassis motors and the arm joint servo-mechanisms, science activity cost values were calculated in terms of required *time* (seconds) and *power* (watts) to execute a traverse to a science target and/or an arm placement. This information was output by the AAPI to the CREST executive for subsequent use by the TVCR.

7. FIELD TRIALS

Figure 14 shows the layout of the nominal scenario used in the field trials, which allowed us to test all aspects of the architecture and the overall system concept. The objective was to determine whether the individual components could function together as a complete system to demonstrate successful closed-loop science autonomy in a representative (ExoMars-like) scenario.

The duration and budget of the project limited the eventual number of trials that could be carried out, thereby limiting the statistical significance of the results for the integrated system; however, the performance of the individual components has also been evaluated in a stand-alone configuration. This allowed more comprehensive testing, e.g., using more complex MER images to evaluate SARA than we could achieve in our test facility.

The planned activity for this simulated sol was to traverse from A to B and carry out periodic assessments of the terrain at discrete waypoints using acquired WAC images. The platform passes a small rock field to its left (near centerline of figure), which is not of science interest. As it nears WP3 there is a set of rocks at 10 o'clock from its position. Some of these are of high scientific value as they exhibit layering features including both horizontal and cross bedding. The challenge therefore was to ignore some rocks, detect those of interest, and escalate the opportunistic analysis to include high-resolution imaging and arm placement.

Figure 15 shows how the nominal plan evolved in the course of the experiment. Images at WP2 were assessed but not deemed to be interesting enough. At WP3 a morphological analysis was carried out on an input WAC image; the resulting analysis showed a high concentration of peak and ridge features, which caused the SARA component to generate a request to take a high-resolution image. This was assessed and authorized by TVCR. The figure also shows the resulting plan deviation where a close-up or level 2 analysis fragment has been inserted into the nominal plan. Once executed, SARA carried out a structural analysis on the high-resolution image.

The output shown in Figure 16(d) shows a significant concentration of horizontal bedding. SARA requested a more detailed analysis of the target via the executive. The executive queried the arm agent to get an approximate measure of the resources required to conduct an arm placement. It then commanded TVCR to parse this request and determine its feasibility. TVCR determined that the request could be

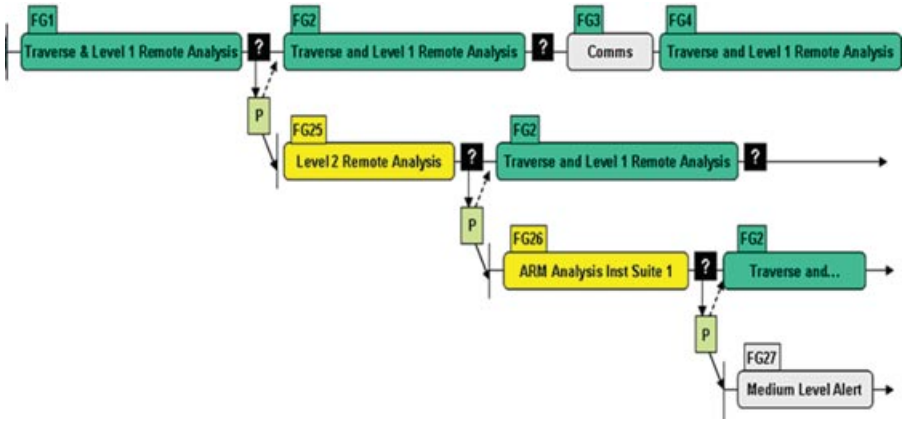

Figure 15. Portion of nominal plan consisting of a set of traverse and analysis (including acquisition of WAC images) tasks on the first line at the top. Requests for the insertion of additional science tasks at WP3 are denoted by a "?" box and TVCR assessment of each request indicated by the "P" box. The continuous line from TVCR shows the outcome, with the dashed line highlighting what would happen if the request was refused, i.e., no replan. Lines 2 and 3 show the result of successful replans with new science tasks (level 2 remote analysis and ARM Analysis Inst Suite 1) inserted by TVCR following its approval of opportunistic requests from SARA (via the executive). Note: the "FG" labels attached to tasks are simply task identifiers.

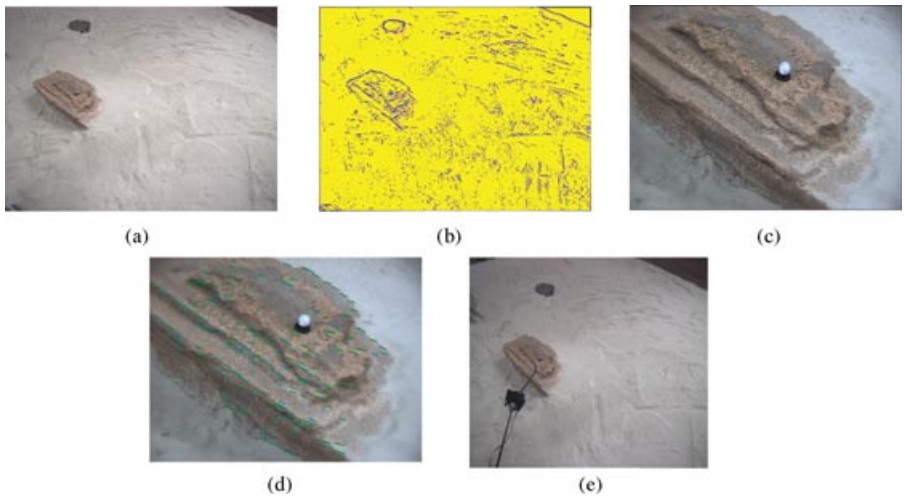

Figure 16. (a) WAC input image. (b) Processed morphology analysis. (c) High-resolution input image. (d) Bedding analysis output. (e) Successful arm placement. Note: Spherule (Vicon marker) for illustration only.

serviced, and the appropriate plan fragment was inserted in the plan. Figure 16(e) shows the resulting arm placement.

Figure 17 shows a graph of our arm placement results from an AAPI calibration trial. The 10 different science target points (i.e., potential SARA input data) were identified, and their actual positions were measured using our Vicon system. Each science target was then processed using the AAPI, and the arm moved to each target position. The actual arm end point was then measured using the Vicon system so that a comparison could be made against each commanded science target position. Prior to the calibration of the AAPI KJO, we experienced considerable

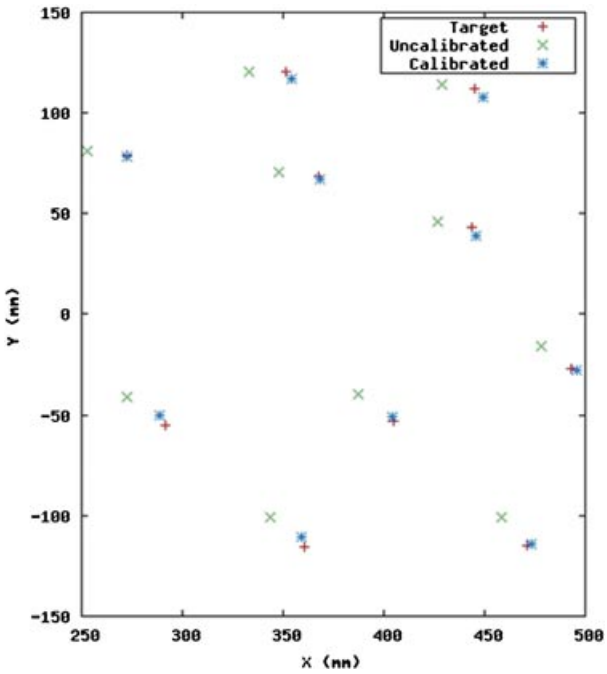

Figure 17. 2D graph of arm placement results: plan view. The 10 different target science points are shown (+) together with the achieved arm placement position for both the calibrated (*) and uncalibrated (×) AAPI KJOs. The uncalibrated points show clearly the placement errors that are due to arm deflection under gravity. The origin for this graph is the mechanical interface between the arm base joint and the rover chassis base plate, i.e., ≈230 mm above the nominal ground plane.

placement errors mainly due to arm deflection; these uncalibrated AAPI points are also shown. Note: As this system did not include an autonomous navigation component, requests for additional analysis were serviced only for rocks within reach of the current position. No mobile target tracking was therefore used or developed. It is assumed that the rover and target have not changed position, therefore, between analysis phases.

It should be noted that the arm kinematics origin was set to be at the mechanical interface between the arm base joint and the rover chassis base plate (i.e., ≈230 mm above the nominal ground plane). Given this origin and adopting conventional robot arm Cartesian axes meant that all of our science target points (for the relatively small rocks that were used) were in the negative Z direction (i.e., lower than the chassis base plate): hence the negative z values in Figures 18 and 19. Analysis of the 10 commanded/actual arm placement results produced a mean arm end point positional accuracy of 3.58 ± 1.793 mm [± 1 standard deviation (SD)].

A week of on-site integration trials at the indoor test facility resulted in the system being able to successfully complete the reference test scenario. As noted earlier, this work focused on the integration of the three main system elements during the limited field trials, so the performance results cannot be assessed in a statistically meaningful way. However, it does give an indication of some of the problems that were encountered and highlights areas that could be improved in future developments. The trials discussed below were performed at the final stages of the integration. Owing to time constraints, trials were on some occasions curtailed (see text below) if a component did not perform as expected in order to address the problem at hand. Given the unavailability of an autonomous navigation component, traverse activities were executed manually by operators waiting on instruction from the execution component trace output. With the exception of the manual drives, SARA, TVCR, AAPI, and the executive operated in an autonomous way starting from the upload of a plan, progressing with telemetry downlinks, and ending with final status reports. Finally, although we experimented with the insertion of microtraverses to reach targets that were close but out of reach in the early stages of integration, these were ruled out in the final trials because of the absence of an autonomous navigation component.

Looking at the performance of the integrated system in more detail, then: SARA operation was relevant with respect to two parts of the terrain, i.e., at WP2, where it had to analyze but ultimately reject targets in view, and at WP3, where it had to correctly identify one rock from four as being the most important and request both close-up analysis and then arm placement activities. As noted in Figure 5, standoff analysis and SVS derivation were based primarily on the structural "roughness" SVS measure, which could be derived from the geomorphological feature vector. For these trials this was set at 0.3 (i.e., ratio of nonplanar terrain that could be considered as rough given occurrence of ridge and channel features, "bedding"). For the last 10 trials, 70% resulted in a successful rejection of the target rocks, i.e., SVS <0.3 for structure. The unsuccessful trials could be attributed to segmentation failure, in which target rocks became fused resulting in an overly high SVS and also scaling was not taken into account and the rover was too close to

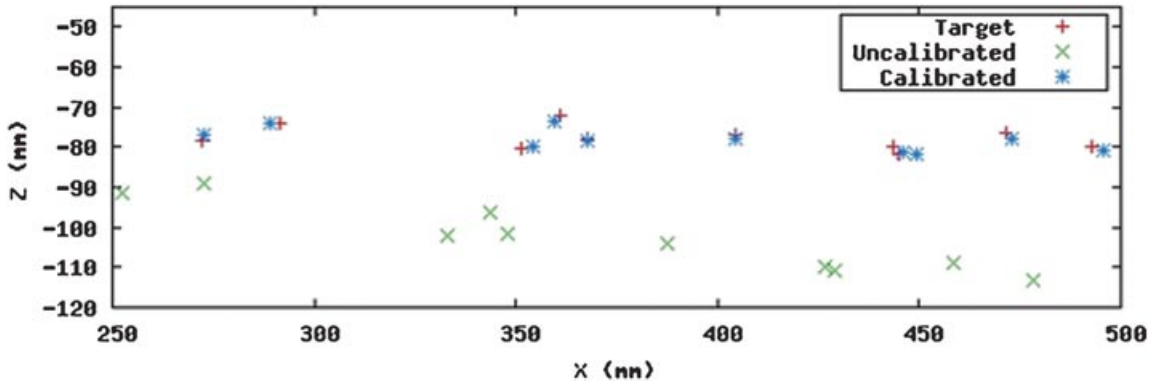

Figure 18. 2D graph of arm placement results: side view.

the target, which resulted in a greater occurrence of ridge and channel features and a false positive from SARA. Between trials the sources of variability were the ambient light conditions, i.e., from artificial light only to a mix of morning/afternoon natural light and artificial lighting, and also small changes in rover positioning created by the manual driving between waypoints. After initial tuning it was found that the relatively small changes in position led to the 30% failure rate noted here.

If SARA behaved incorrectly at WP2, the trial run was aborted because of time constraints. For the seven trials that were successful, the trials proceeded and SARA standoff analysis at the true target field at WP3 was successful for five of the seven trials. The failed runs were again due to poor segmentation or an inability to detect the effect of scale on the results. In both cases the neighboring rock was identified as the target of interest, and although this rock was of the same bedding type, it was out of reach of the instruments and would have resulted in a poor use of resources for close-up analysis in a real mission. For the small number of successful standoff analyses at WP3, the close-up analysis correctly detected a sufficient degree of structural bedding. The target SVS threshold was set at 90 based primarily on the identification of dense cross bedding made up of individual lines, which were >9 pixels long. The average SVS was 140 with a ±20 spread for those runs. For each successful SARA standoff, TVCR authorized close-up analysis and ARM placements by performing a successful replan. For those ARM placements that were invoked after successful close-up analysis, the ARM movements were made as requested with an accuracy that was commensurate with that outlined in Figure 17.

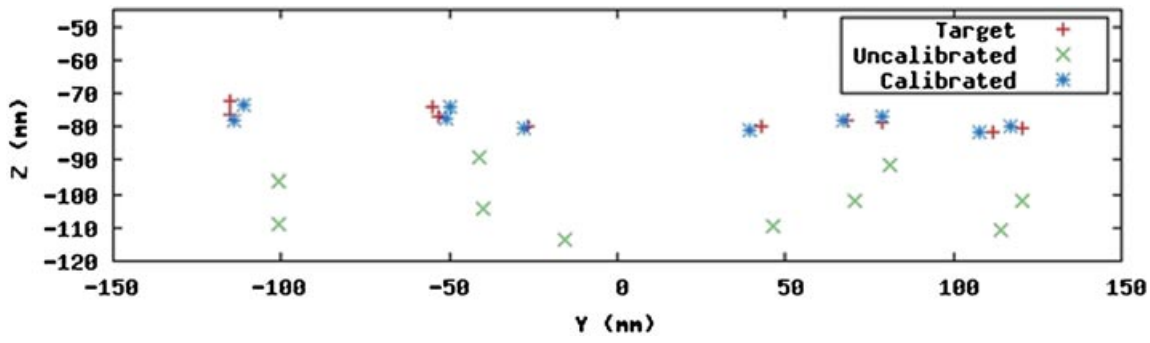

Figure 19. 2D graph of arm placement results: front view.

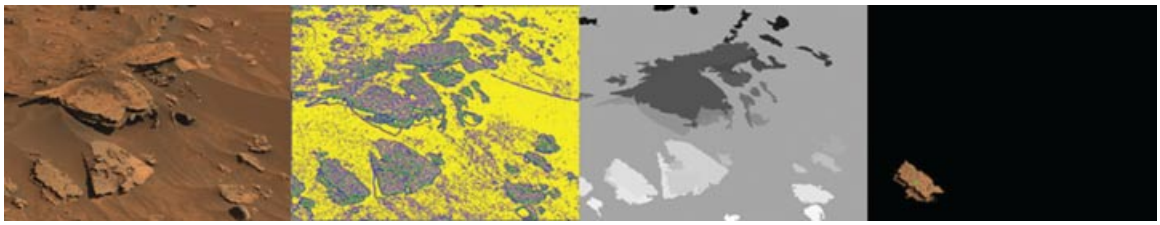

Figure 20. MER Mars test data. Passing this image through the SARA obtains a target on the right-hand image. This was derived from using the feature (e.g., ridge, channel, or peak) density obtained from the second image and the segmented region mask from the third. Although real imagery was used, the system is assisted by the color difference between the soil and rocks.

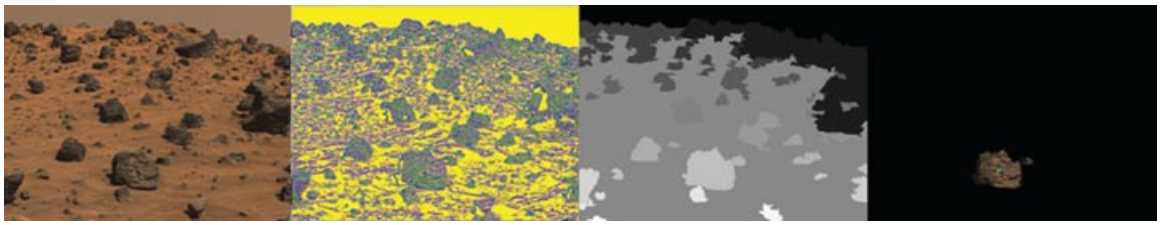

Figure 21. MER Mars test data. Again a target was selected in the right-hand image, which was derived using the feature density characteristics and the segmentation of heterogeneous regions. Again it can be argued that the system was helped by the contrast between the soil and rocks.

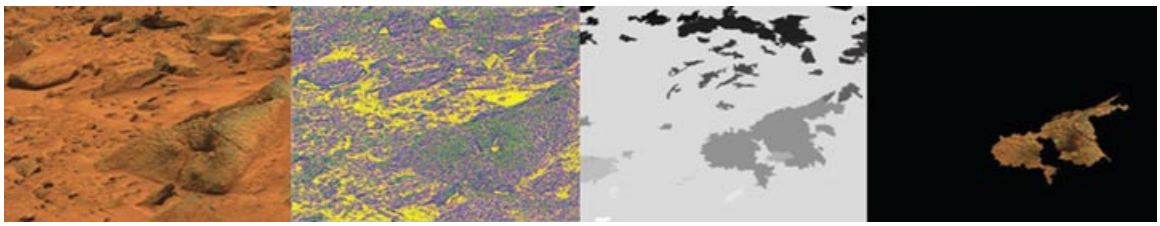

Figure 22. MER Mars test data. This sequence of images shows that when the rocks are very similar to their backgrounds, it becomes hard to distinguish them. Although a target was found, its boundaries are not correct. This could be addressed by changing the contextual model and parameter configuration file, providing better values that will decrease the expected contrast difference.

8. MARTIAN TEST DATA

To test the feasibility of the SARA system and the initial target extraction methods, actual Martian data were obtained from the MER[2] panoramic image database. Real data were used for testing due to the excessive contrast between background terrain and targets of interest in our simulated Mars environment used during the integration experiments. The main differences between the real and experimental setup are the contrast between the rocks and the surrounding terrain and sky, etc., and the density, shape, size, and distribution of the rocks and other interesting features. Although the SARA system is in the early stages of its development and is limited in actual analysis capability, the results shown in Figures 20–25 are with no reconfiguration of system parameters as used during the field trials. Figures 20–23 show how SARA performed with MER, standoff scale images. Figures 23 and 24 show close-up performance.

[2]http://www.nasa.gov/mission_pages/mer/index.html

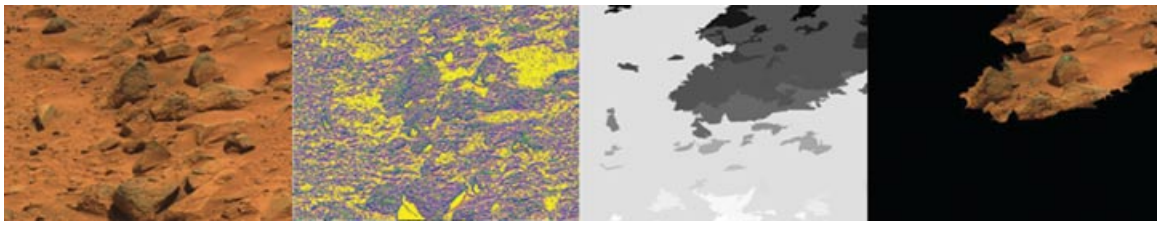

Figure 23. MER Mars test data. Although this image is similar to Figure 22, the segmentation algorithm could not distinguish the region at the top right (third image), which leads to the result in the right-hand image, where the target is not a single rock but several, and it is coincidental that the target position is actually on the rock. Again this could be addressed through the contextual model and configuration files, as well as an adaptive segmentation algorithm.

The results shown in Figures 20–23 show that the geomorphology analysis approach for standoff analysis performs reasonably well with complex 2D MER images. This approach, which was developed and tested in our aerobot studies (Woods et al., 2008), works best on 3D data such as DEMs, and we anticipate improvements in robustness if DEMs were also assessed. There is of course an additional computational cost to consider for 3D data analysis. This implementation selected targets on the basis of absolute feature density. In the aerobot work we used a fuzzy logic–based system to apply a more sophisticated analysis to raw feature data in order to extract targets with particular characteristics.

To complete the SARA analysis test using real-world data, high-resolution images from the MER mission were also used. An important note here is that presently SARA looks only for bedding and limited textual information. In normal operations the SARA would have requested a high-resolution close-up image (from the AAPI) of the target from the standoff analysis.

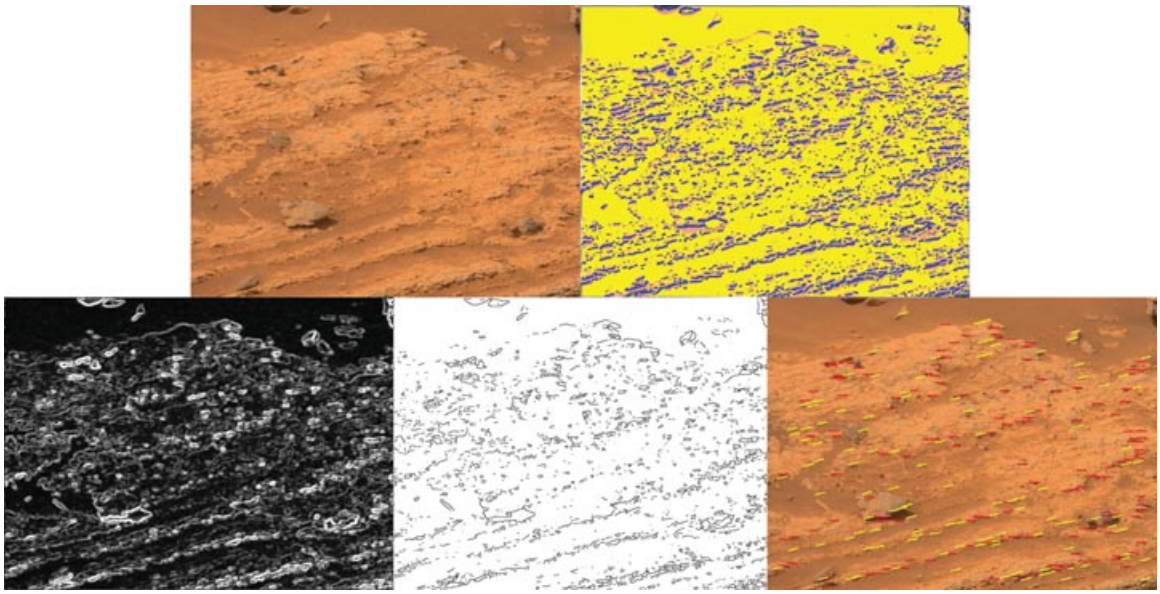

Figure 24. MER Mars close-up image test data. Features are detected (top right); this describes surface texture and roughness. Sobel edge detection and morphological gradient combination (bottom left) allows the detection of the changing boundaries, which are then filtered and thinned (bottom center). A line mask is then passed over the thinned image, identifying potential lines and their directions and their lengths. This classification is superimposed onto the original image (bottom right). Bottom right image shows the identified bedding having an orientation of between 0 and 14 deg approximately with respect to the x axis (horizontal) of the image.

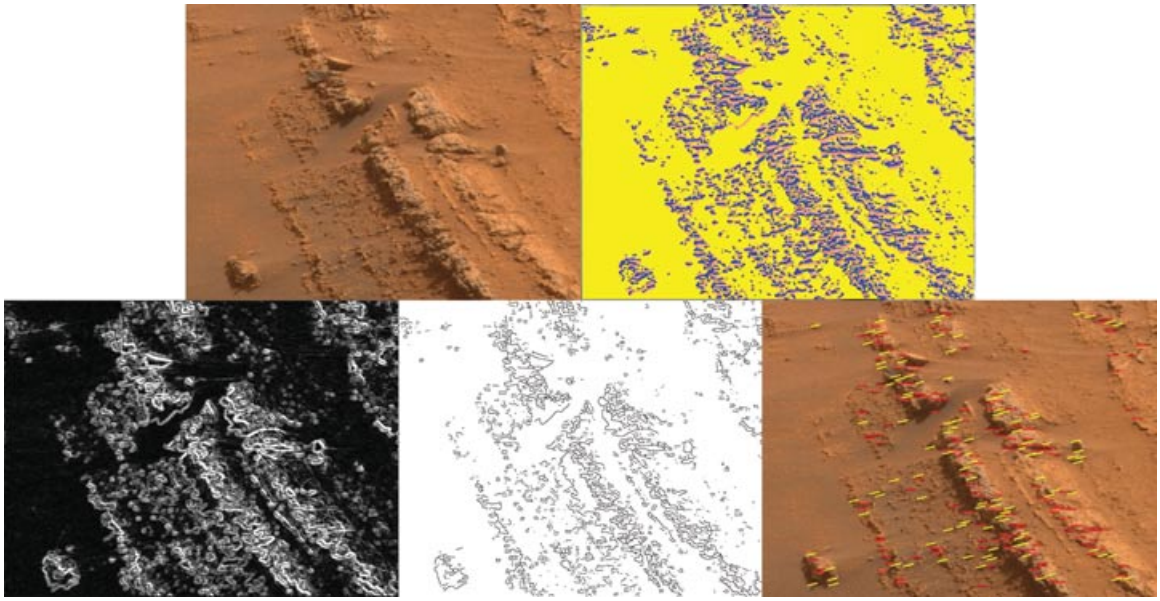

Figure 25. Similar to Figure 24, but this time the bedding analysis is incorrect as horizontal lines are again identified. This is a result of the scale of the image and the line mask as there are horizontal lines present in the image from wind deposition and shadows from the sun position.

Figures 24 and 25 again show that the initial results from the SARA at processing actual Martian data are promising. To improve the analysis and classification performance further, information regarding scale and 3D profiles would be very useful as these would reduce the errors from using the masks. This information would also help in the classification of texture, as the scale and shape are very important. The 3D information would also help eliminate the effect of shadowing on the system.

9. TVCR PLAN MODIFICATION EVALUATION

TVCR is built on a plan validation system, implemented in VAL (Howey, Long, & Fox, 2004). The performance of TVCR is dominated by a series of attempts to validate plans using VAL, with plan modifications being achieved by a series of edits that are subjected to validation before being released. The edits themselves are cheap, because they amount simply to confirming that the conditions hold in order to allow the addition of predefined plan segments, or else to adjust plans as elements are removed. In both cases, the numbers of constraints that must be considered is small, because plan fragments generally have local relationships within the plan structure. Thus, the most important function in determining the execution costs for TVCR is the validation task. To give an indication of the cost of plan validation, we have run the validator on a collection of competition plans from the 3rd International Planning Competition (Long & Fox, 2003). The test was executed on a desktop PC running Linux with a 2.6-GHz processor and 500 MB of memory. The collection contains 5,381 plans, with a combined length of more than 185,000 steps. This complete collection requires 14 min and 14 s (elapsed time) to validate (not all, but most plans are valid). This is equivalent to an average of 160 ms to validate each plan, with an average of more than 34 steps per plan. Note that this measurement includes separate invocation of the validator from the command line for each plan.

A separate test in which multiple plans for the same problem were validated in a single batch (so only one invocation of the validator was required) completed a validation of 20 plans, averaging 364 steps each, in just under a second. This shows that without the overhead of starting the validator and reading in the domain and problem files, the average time to validate a plan is as little as 50 ms.

In a plan modification cycle, the maximum number of validation calls is determined by the number of plan fragments the plan contains and how many might have to be removed in order to arrive at a plan

that fits the resource envelope for the execution cycle. In practice, this is typically limited to about 10 calls. In fact, far fewer are likely to be required: in our tests, only four or five calls were needed. Thus, the total resource requirement for this system is less than 1 s of CPU time. Of course, onboard CPU resources are likely to be significantly less powerful than the desktop used for evaluation (in a space science environment). A space-qualified CPU runs at 100 MHz, so we might hope to lose two orders of magnitude in the transition, suggesting that the upper bound on the required CPU time would be about 2 min, but less than half of this would be required on typical runs.

It is very difficult to give a thorough empirical evaluation of the plan modification strategy itself, because this depends on the properties of the fragments that are supplied to encode opportunities and the availability of resources with which to exploit the opportunities. The approach TVCR adopts is conservative and greedy, so it will attempt to maximize the number of high-priority fragments it fits into the time line, but with minimal alteration of the existing time line and always subject to the constraints on use of fragments and on validity of the time line. This means that TVCR will not fail to produce a valid time line, but it might fail to find an optimal ordering of fragments to achieve the maximal operational use of resources. We did not experience any such examples, but it is not hard to construct artificial problems with this property. However, TVCR does not produce bad solutions in these cases, merely suboptimal solutions (according to the maximum science return criterion). We believe that solutions that push the envelope of resource availability to achieve maximal returns are likely to be more fragile and prone to failure, so a solution that appears optimal from the perspective of predicted science return might well be suboptimal from the perspective of actual science return.

10. LESSONS LEARNED

Owing to the short duration of the CREST project (1 year), only very basic criteria were used for scientific assessment (i.e., presence of layering, absence of layering, or albedo from gray scale). Furthermore, the need to incorporate the planning agent and other system architecture components within the development and testing limited the time available for generic science evaluation. Clearly, this was necessary to explore and assess the feasibility of the proposed the architectural approach. It suggests that for future work, the science evaluation component could be developed independently. Integration could then be repeated when the component is more mature.

Although the identification methods used proved reasonably robust with respect to the image data obtained during the test campaigns, differences in the certainty of the analysis results were obtained under various lighting conditions. These differences were primarily due to the lighting intensity and direction of shadowing effects caused by either changes in the ambient Mars Yard environment or changes in the position of the platform between groups of trials. On occasion, these variations were quite subtle and resulted in the introduction of extra information on target objects. These effects could also obscure target features to some extent. All of these issues exposed the well-established problem of threshold sensitivity in computer vision applications.

The use of 3D information generated by the pancam/navcam systems would greatly improve the geological identification and classification process. The additional perspective will allow easier extraction of the granule size and shapes and the variation in the depths of the bedding structures and coarseness of the surfaces, similar to field geologists when they are interacting with their environment. The perspective on an image can also influence the result of the processing as the angle from which a rock is presented to the camera can mean that interesting features are out of view.

As far as planning and plan modification are concerned, we found that modeling the domain in PDDL was straightforward and gave access to a considerable collection of useful tools. However, PDDL does not include any way to capture the organizational structure of plans or parts of plans. The expression of plan constraints and plan fragments is an additional layer built over the PDDL domain description. It is interesting to observe that a similar constraint language was found necessary to support planning in MAPGEN (Bresina & Morris, 2006), and many of the observations made in the use of MAPGEN, although it is a ground-based tool for mixed-initiative planning, apply equally to the TVCR onboard planning context. The challenges of restricting the opportunistic extension or modification of plans, within limited computational resources, remain significant. Experiences with the Autonomous Spacecraft Experiment on Earth-Observing 1 (Chien et al., 2005), although grounded in a more extensive experience with deployment, show many similarities with our own.

A number of lessons have been learned from the AAPI work. First, the mechanics of the arm and pan and tilt mechanism used during the field trials must be improved. Expediency drove us to use a simple three-DoF arm and to use commercial off-the-shelf (COTS) servo mechanisms to drive the arm and pan/tilt rotation axes. Although this provided a simple and relatively quick solution, it did generate pan-cam pointing and arm accuracy placement problems. Furthermore, low-mass carbon fiber structures were used for the mast structure and arm link lengths. Although this was acceptable for the pan/tilt mechanism, it did generate arm deflections (as shown by the measured KJO values required to compensate for this) despite the fact that there was very little mass at the arm end point. Although low-mass structures are essential for flight hardware (although carbon fiber would not be appropriate for an exobiology mission), ultra-high-quality, zero-backlash gearing must be used, and any kinematics model must incorporate an arm deflection model (both bending and torsion effects). This is an area that is being worked on as part of our current AAPI to Beagle 2 Development Model (DM) arm porting activity. The Beagle 2 arm has five DoFs and has been studied as a possible baseline arm for the ExoMars mission.

Second, calibration is a vital area that must be given a great deal of thought to plan and execute. All kinematics devices must be calibrated together with any camera intrinsic and extrinsic properties. For the work that we are performing, then, arm/camera cross calibration is also required. Calibration is a time-consuming activity and shortcuts are not advisable. Although we have considerable experience with arm calibration (having calibrated the Beagle 2 Development Model and Flight Model arm) and camera calibration (having been team members of the Beagle 2 Stereo Camera System), tracking down every error source, systematic or otherwise, is a major activity. Given the mechanical inadequacies of our arm and pan/tilt mechanics, we did manage to achieve a mean arm end point positional accuracy of 3.58 ± 1.793 mm (± 1 SD).

It should be noted that the current requirements for the ExoMars mission dictate a contact instrument positional accuracy better than 5 mm (including ± 1 SD) and an instrument approach direction accuracy of less than 0.5 deg (relative to the desired approach trajectory, and including ± 1 SD). However, it must be emphasized that these final error limits are after, first, a stereo camera capture and 3D target processing, second, a subsequent 3-m traverse to the vicinity of the science target, and finally, the required arm motions and instrument placement activities. Automation of such an activity sequence while keeping within the required resultant positional and angular error limits will be not be easy. Owing to the contact instrument positional errors that are inherent to any arm placement strategy, it is highly likely that an approach similar to that used for Beagle 2 will be adopted for ExoMars. This involved the use of a force sensor at the mechanical interface between the instrument head (PAW) and the robot arm. Hence the final approach to any science target contact incorporated a *do move while force sensor* $\leq$ *preset threshold* joint control strategy.

Third, although we found that the stereo matching and occlusion detection algorithm that we adopted for our field trials performed well with respect to high-disparity map accuracy, we did find that well-informed algorithm initialization parameters were required. Although this would be acceptable for a Mars mission that had a high degree of ground-based intervention, future missions requiring a high degree of rover autonomy would find that a different algorithmic approach is necessary. Furthermore, adopting an algorithm that has been developed for standard desktop computing applications would most likely completely overpower the relatively meager computing resources that are available to the current state-of-the-art planetary rover. Space qualification is a costly and time-consuming activity, and although progress is being made, it will be some time before we come close to the desktop computing resources that we all take for granted. Until we reach this situation, then, onboard planetary rover stereo image processing will require "smart" but computationally minimal algorithms. We refer to such algorithms as "simple smarts," and we believe this to be an overlooked research area.

Fourth, we do believe that there is an argument for investigating the use of FPGA technology for many of the AAPI algorithms used in this study. Certainly elements of the WAC image processing, the stereo triangulation, the forward and inverse kinematics, and the various transforms developed as part of the AAPI process are of sufficient maturity to warrant an FPGA approach. We envisage a dedicated arm agent and a perception agent whereby the processing is not reliant on the planetary rover core processor. Certainly the arm joint motor control needs close coupling between speed commanding and

motion sensing, and an FPGA would be ideal for such activities. If we have to wait for improved space-qualified computing hardware, then an obvious approach is to have a multiprocessor solution (interconnected using SpaceWire[3] or a multicore system on a chip solution). FPGAs form an ideal component technology within such a distributed computing architecture.

Finally, our least-squares Helmert transform–based approach used to map candidate science target 3D data from camera Cartesian space to robot arm kinematics Cartesian space has shown that in situ camera/arm calibration is possible. Even after extensive calibration activities, one must realize that all planetary rover components must be sterilized and baked at 125°C for 48 h (or use peroxide-plasma methods). It must then withstand the vibrations of launch, 20–2,000 Hz, +9 to −9 dB/oct, the vacuum, and −100°C temperatures of space. Upon impact with the surface of Mars, even with the protection of gasbags, the rover could experience shock forces as high as 200× gravity. Once deployed it must operate within an atmosphere that has a breakdown voltage as low as 100 V, very fine blown dust, and a diurnal temperature range of −100 to +30°C. This all goes to emphasize that even with a pristine calibrated robot prior to launch, one cannot assume that it will arrive on the surface of Mars with all of its calibration parameters intact. In situ (on Mars) calibration (correction) procedures need to be developed in parallel with the conventional prelaunch procedures. We believe that closed-loop camera processing and arm placement control methods are a fertile area for future research.

We learned a number of lessons from how the trials themselves were conducted. First, one of the main features of this work was integration of individually complex components. Several sustained (i.e., full team complement over a number of days), integration testing and preparation sessions were essential in the run-up to the final field trials. Second, it is important to explicitly define and specify the test trial trajectories in situ and on paper before the end of preparation sessions so that a common understanding can be truly realized. It is easy to verbally affirm a common understanding during preparation in the vicinity of the test area when in fact subtly different interpretations may be formed. If these are not checked, effort can be misplaced in the off-site individual follow-up work. Finally, if the final field trials are temporally constrained (e.g., access to the test site resource being limited), then it is important that the trials are actively and tactically managed to ensure that the main goals are achieved. This requires prioritization during problem resolution and initiation of appropriate parallel activities when bottlenecks occur.

[3]SpaceWire is an asynchronous spacecraft communication-switched network based in part on the IEEE 1355 standard of communication. It is coordinated by ESA in collaboration with several European space companies and other space agencies (NASA and JAXA).

11. CONCLUSIONS

Implementing science autonomy for a robotic exploration platform is a challenging but attractive goal. It requires a comprehensive system design that coordinates and combines a number of complex components. This work has shown that such a system concept can be developed for a mission such as ExoMars, which evolves the current operations paradigm in a practical way. Three generic elements are required to build this system. First, there is science assessment and response. We have outlined a framework for the development of an approach to autonomous field science–based analysis of fundamental geological attributes. The SARA implementation demonstrated here has addressed only the basic elements of this framework, and clearly the evolution of this capability is a long-term prospect. However, we believe that it can be deployed in an incremental way as the concept evolves. The next step would be to develop this element further in a stand-alone form with which to test the proposed methodology more thoroughly. Such a model would be independent of any system architecture or planning software and concentrate on the generic concepts of science evaluation and assessment. At the core of the model would be a relational database populated with a broad and diverse library of structural, textural, and compositional reference data ranging from simple gray scale cartoon images through to multispectral image cubes and compositional spectra. In addition to the scientific content, the database would also include engineering and environmental information relevant to the mission, instrumentation, and planetary body. Suitable data have already been accumulated from existing missions and analog studies (Pullan et al., 2008).

The second critical component is a plan validation and replanning element. This study has

demonstrated the suitability of the TVCR concept for this particular task. Such a component may be used on ExoMars to support time line optimization. It is beneficial to know, therefore, that it could also support autonomous science should it be deployed. The third generic component is the autonomous control of active platform elements such as a robotic arm or locomotion system. In this work we have demonstrated a concept for autonomous approach and placement of a robotic arm.

The arm used for the field trials was a three-DoF device, and simple joint-by-joint control methods were used. Currently we are working on a five-DoF arm, namely the Beagle 2 development model arm. Here we have developed full joint-interpolated-motion with simultaneous joint motor speed control. Arm end-effector placement trajectories can be defined for this five-DoF arm, and the generated science target rock DEM data can be used to calculate a rock surface normal for arm placement trajectory determination. As we have detailed CAD models of the Beagle 2 arm and associated mounting structure, we propose to use a bounding sphere approach to check for potential arm collision situations. Currently we are porting our AAPI software to the Beagle 2 arm application, and the results from this work will be reported in future literature.

The integration of an autonomous science assessment with both planning *and* autonomous arm placement has been a unique contribution of this work. The main philosophical contribution has been our attempt to provide an implementation of the way geologists hierarchically explore scenes of interest over a range of scales in a continuous manner: i.e., acquire, assess, decide what data and at what scale would help explain the target, acquire, and repeat, etc. This has been achieved by providing a framework to allow different types of autonomy (science assessment and response, planning, approach, and placement) to collaborate interactively to pursue scene investigation over a number of iterative steps.

In conclusion we believe that, although challenging, science autonomy in some form will be essential for future exploration missions given the inefficiency of using solely Earth-based science analysis and planning.

The need to optimize the collection, selection, and transmission of science data could greatly benefit from the use of such technology and may ultimately be required to ensure that nominal science goals are achieved.

ACKNOWLEDGMENTS

We would like to extend our gratitude to the U.K. STFC represented by Mrs. Sue Horne, who funded this work as part of the U.K. CREST program, and to Dr. Laurence Tyler, IMAP, Aberystwyth University, for his contribution to the AAPI calibration work. For more information and videos go to http://www.scisys.co.uk/casestudies/space/cs_robotsci.asp.

REFERENCES

Backes, P., Diaz-Calderon, A., Robinson, M., Bajracharya, M., & Helmick, D. (2005, March). Automated rover positioning and instrument placement. In IEEE Aerospace Conference, Big Sky, MT (pp. 60–71).

Bouguet, J.-Y. (2008). Camera calibration toolbox for Matlab®. Supplied by Computer Vision Research Group, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA.

Bresina, J. L., Golden, K., Smith, D. E., & Washington, R. (1999, June). Increased flexibility and robustness of Mars rovers. In Proceedings of Fifth International Symposium on Artificial Intelligence, Robotics and Automation in Space, ESTEC, Noordwijk, The Netherlands (p. 167).

Bresina, J. L., & Morris, P. H. (2006, July). Mission operations planning: Beyond MAPGEN. In Proceedings of the 2nd IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT'06), Pasadena, CA (pp. 477–484).

Castaño, R., Estlin, T., Anderson, R. C., Gaines, D., Castaño, A., Bornstein, B., Chouinard, C., & Judd, M. (2007). OASIS: Onboard autonomous science investigation system for opportunistic rover science. Journal of Field Robotics, 24(5), 379–397.

Castano, R., Estlin, T., Gaines, D., Castano A., Bornstein, B., Chouinard, C., Anderson, R. C., & Judd, M. (2006, March). Automated target selection for opportunistic rover science. In Proceedings of Lunar and Planetary Science, XXXVII, League City, TX (abstract 2434).

Chien, S. A., Sherwood, R., Tran, D., Cichy, B., Rabideau, G., Castaño, R., Davies, A., Mandl, D., Frye, S., Trout, B., D'Agnostino, J., Shulman, S., Boyer, D., Hayden, S. C., Sweet, A., & Chista, S. (2005, July). Lessons learned from autonomous spacecraft experiment. In Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), Industrial Applications Track, Utrecht, The Netherlands (pp. 11–18).

Chien, S., Smith, B., Rabideau, G., Muscettola, N., & Rajan, K. (1998). Automated planning and scheduling for goal-based autonomous spacecraft. IEEE Intelligent Systems, 13(5), 50–55.

Erol, K., Nau, D., & Hendler, J. (1994, July). HTN planning: Complexity and expressivity. In Proceedings of the National Conference on Artificial Intelligence (AAAI-94), Seattle, WA (pp. 1123–1128).

Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. International Journal of Computer Vision, 59(2), 167–181.

Fox, M., Howey, R., & Long, D. (2005, July). Validating plans in the context of processes and exogenous events. In Proceedings of the 20th National Conference on Artificial Intelligence (AAAI-05), Pittsburgh, PA (pp. 1151–1156).

Fox, M., & Long, D. (2003). PDDL2.1: An extension to PDDL for expressing temporal planning domains. Journal of AI Research, 20, 61–124.

Gazis, P. R., & Roush, T. (2001). Autonomous identification of carbonates using near-IR reflectance spectra during the February 1999 Marsokhod field tests. Journal of Geophysical Research, 106, 7765–7773.

Gonzalez, R. C., & Woods, R. E. (1992). Digital image processing (pp. 518–560). Upper Saddle River: Addison-Wesley Publishing Company, Inc.

Gulick, V., Morris, R. L., Ruzon, M. A., & Roush, T. L. (2001). Autonomous image analysis during the 1999 Marsokhod rover field test. Journal of Geophysical Research, 106, 7745–7763.

Howey, R., Long, D., & Fox, M. (2004, November). VAL: Automatic plan validation, continuous effects and mixed initiative planning using PDDL. In Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence, Boca Raton, FL (pp. 294–301).

Huntsberger, T., Cheng, Y., Stroupe, A., & Aghazarian, H. (2005, August). Closed loop control for autonomous approach and placement of science instruments by planetary rovers. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005 (IROS 2005), Edmonton, AB, Canada (pp. 3783–3790).

Ingham, M., Ragno, R., & Williams, B. C. (2001, June). A reactive model-based programming language for robotic space explorers. In Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space, St. Hubert, Canada.

Long, D., & Fox, M. (2003). The 3rd International Planning Competition: Results and analysis. Journal of AI Research, 20, 1–59.

Nebel, B., & Koehler, J. (1995). Plan reuse versus plan generation: A theoretical and empirical analysis. Artificial Intelligence, 76(1–2), 427–454.

Pedersen, L. (2000, November). Robotic rock classification and autonomous exploration. Doctoral dissertation (Tech. Rep. CMU-RI-TR-01-14). Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Pedersen, L., Deans, M., Lees, D., Rajagoplan, S., & Smith, D. (2005, September). Multiple target single cycle instrument placement. In Proceedings of the 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Munich, Germany.

Pedersen, L., Kunz, C., Sargent, R., Madison, R., Backes, P., Bajracharya, M., Clouse, D., & Nesnas, I. (2006, September). Performance evaluation of handoff for instrument placement. In AIAA Space 2006, San Jose, CA (pp. 1–19).

Pullan, D., Westall, F., Hofmann, B. A., Parnell, J., Cockell, C. S., Edwards, H. G., Villar, S. E., Schröder, C., Cressey, G., Marinangeli, L., Richter, L., & Klingelhöfer, G. (2008). Identification of morphological biosignatures in Martian analog field specimens using in situ planetary instrumentation. Astrobiology, 8, 119–156.

Rabideau, G., Knight, R., Chien, S., Fukunaga, A., & Govindjee, A. (1999, June). Iterative repair planning for spacecraft operations using the ASPEN system. In Artificial Intelligence, Robotics and Automation in Space, Proceedings of the Fifth International Symposium, ISAIRAS '99, ESTEC, Noordwijk, The Netherlands (pp. 99–106).

Shaw, A., & Barnes, D. (2003, October). Landmark recognition for localisation and navigation of aerial vehicles. In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2003), Las Vegas, NV (pp. 42–47).

Sherwood, R., Tran, D., Cichy, B., Rabideau, G., Castano, R., Davis, A., Mandl, D., Trout, B., Shulman, S., Boyer, D., & Chien, S. (2005). Using autonomy flight software to improve science return on Earth observing one. Journal of Aerospace Computing, Information, and Communication, 2(4), 196–216.

Simmons, R., & Apfelbaum, D. (1998, October). A task description language for robot control. In Proceedings of the Conference on Intelligent Robotics and Systems, Victoria, BC, Canada (pp. 1931–1937).

Stow, D. A. V. (2003). Sedimentary rocks in the field. London: Manson Publishing.

Taylor, R. H. (1979). Planning and execution of straight line manipulator trajectories. IBM Journal of Research and Development, 23(4), 424–436.

Thompson, D., & Castano, R. (2007, March). Performance comparison of rock detection algorithms for autonomous planetary geology. In Proceedings of 2007 IEEE Aerospace Conference, Big Sky, MT (pp. 1–9).

Thompson, D., Smith, T., & Wettergreen, D. (2005, September). Data mining during rover traverse: From images to geologic signatures. In Proceedings of 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Munich, Germany.

Thompson, D. R., Smith, T., & Wettergreen, D. (2006, March). Autonomous detection of novel biologic and geologic features in Atacama Desert rover imagery. In LPSC XXXVII, Houston, TX (abstract 2085).

Van Winnendael, M., Baglioni, P., & Vago, J. (2005, September). Development of the ESA ExoMars Rover. In Proceedings of 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Munich, Germany.

Wettergreen, D., Cabrol, N., Heys, S., Jonak, D., Pane, D., Smith, M., Teza, J., Tompkins, P., Villa, D., Williams, C., Wagner, M., Waggoner, A., Weinstein, S., & Whittaker. W. (2005, August). Second experiments in the robotic investigation of life in the Atacama Desert of Chile. In Proceedings of 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Munich, Germany.

Wood, J. D. (1996). The geomorphological characterisation of digital elevation models. Ph.D. thesis, University of Leicester, Leicester, U.K.

Woods, M., Baldwin, L., Wilson, G., Hall, S., Pidgeon, A., Long, D., Fox, M., Aylett, R., & Vituli, R. (2006, June). MMOPS: Assessing the impact of on-board autonomy for deep space missions. In Proceedings of SpaceOps, Rome, Italy (AIAA-2006-5977).

Woods, M., Shaw, A., Evans M., Barnes, D., Summers, P., Pullan, D., Paar, G., & Bauer, A. (2008, February). Image based localisation and autonomous image assessment for a Martian aerobot. Proceedings of 9th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Hollywood, CA.

Zhang, Z. (2000). A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), 1330–1334.

Zitnick, C. L., & Kanade, T. (2000). A cooperative algorithm for stereo matching and occlusion detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(7), 675–684.